

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:32:18 / Search time 31.6461 Seconds

(without alignments)
2832.275 Million cell updates/sec

Title: US-09-856-061-2
Perfect score: 2316
Sequence: 1 MTSQGNKRTTKEGFDLRFQ.....QPLPLARLLTQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_minc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	11 Q9JMJ3	Q9JMJ3 mus musculus
2	2305	99.5	435	11 Q9QZP2	Q9QZP2 mus musculus
3	1132	48.9	376	4 Q9P2U9	Q9P2U9 homo sapien
4	410.5	17.7	533	11 Q922M0	Q922M0 mus musculus
5	410	17.7	525	11 Q920L0	Q920L0 ratius norv
6	381	16.5	530	13 Q9DC07	Q9DC07 gallus gall
7	319	13.8	456	4 Q8WV28	Q8WV28 homo sapien
8	319	13.8	456	4 Q75498	Q75498 homo sapien
9	315.5	13.6	552	13 Q9YGC1	Q9YGC1 gallus gall
10	309.5	13.4	433	4 Q75499	Q75499 homo sapien
11	297.5	12.8	457	11 Q9QJN3	Q9QJN3 mus musculus
12	287.5	12.4	457	11 Q88504	Q88504 mus musculus
13	219.5	9.5	297	11 Q9DA13	Q9DA13 mus musculus
14	188.5	8.1	516	11 Q54737	Q54737 mus musculus
15	167.5	7.2	471	4 Q9UCX5	Q9UCX5 homo sapien
16	165.5	7.1	594	4 Q92529	Q92529 homo sapien

17	165.5	7.1	594	4 Q8TPA2	Q8TPA2 homo sapien
18	150.5	6.5	486	4 Q9UCX4	Q9UCX4 homo sapien
19	150.5	6.5	812	15 Q85466	Q85466 773 sarcoma
20	143.5	6.2	662	5 Q9N3S5	Q9N3S5 caenorhabdi
21	141.5	6.1	474	11 Q61120	Q61120 mus musculus
22	141	6.1	728	4 Q9UPH9	Q9UPH9 homo sapien
23	139.5	6.0	286	5 Q95PX0	Q95PX0 caenorhabdi
24	139.5	6.0	398	5 Q95PW9	Q95PW9 caenorhabdi
25	138.5	6.0	335	4 Q96CZ8	Q96CZ8 homo sapien
26	138.5	6.0	926	4 Q9HOK1	Q9HOK1 homo sapien
27	137.5	5.9	594	11 Q70143	Q70143 ratius norv
28	136.5	5.9	477	11 Q70142	Q70142 ratius ralt
29	134	5.8	309	11 Q9QVZ0	Q9QVZ0 mus sp. shb
30	134	5.8	461	11 Q63789	Q63789 ratius norv
31	134	5.8	1599	5 Q18892	Q18892 caenorhabdi
32	133.5	5.8	217	11 Q9CX39	Q9CX39 mus musculus
33	132.5	5.7	904	5 Q9VMJ4	Q9VMJ4 dirosophila
34	132.5	5.7	944	5 Q8T9B3	Q8T9B3 drosophila
35	130.5	5.6	722	13 Q8U0U2	Q8U0U2 xenopus lae
36	129	5.6	559	11 Q91252	Q91252 mus musculus
37	129	5.6	645	6 Q95JY5	Q95JY5 macaca fasc
38	128.5	5.5	112	11 P70305	P70305 mus musculus
39	128.5	5.5	393	2 Q93CA3	Q93CA3 treponema p
40	128.5	5.5	854	11 Q9QZS8	Q9QZS8 mus musculus
41	128.5	5.5	1430	11 Q8VHK2	Q8VHK2 ratius norv
42	128	5.5	623	6 Q95JL7	Q95JL7 macaca fasc
43	128	5.5	651	11 P97504	P97504 mus musculus
44	128	5.5	1160	11 Q9ESC8	Q9ESC8 mus musculus
45	128	5.5	4498	13 Q93291	Q93291 fugu rubrip

ALIGNMENTS

RESULT 1	
Q9JMJ3	PRELIMINARY: PRT; 435 AA.
AC Q9JMJ3;	
DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE MIST.	
GN CLNK OR MIST.	
OS Mus musculus (Mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX NCBI_TaxID=10090;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=20208989; PubMed=10744659;	
RA Gotsuka R., Kanazashi H., Sasanna H., Fujimura Y., Hidaka Y.,	
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.,	
RT "A BASH/SLP-76-related adaptor protein MIST/Clnk involved in Ige	
RT receptor-mediated mast cell degranulation.";	
RL Int. Immunol. 12:573-580(2000).	
DR EMBL: AB021220; BAA96240.1; --	
DR HSSP: P23727; 2PNB.	
DR MGD: MGI:1351468; Clnk.	
DR InterPro: IPR000980; SH2.	
DR Pfam: PF00017; SH2; 1.	
DR PRINTS: PR00401; SH2DOMAIN.	
DR Prodom: PD000093; SH2; 1.	
DR SMART: SM00252; SH2; 1.	
DR PROSITE: PS00001; SH2; 1.	
SQ SEQUENCE 435 AA; 49513 MW; 4D1BD3E2F0C61ED6 CRC64;	
Query Match	100.0%; Score 2316; DB 11; Length 435;
Best Local Similarity	100.0%; Pred. No. 8.3e-180;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY 1 MTSQGNKRTTKEGFDLRFQNVSLKNRSPSLSSAKGRCRAVLEPLPHRRRLAGVPGG 60	
DB 1 MTSQGNKRTTKEGFDLRFQNVSLKNRSPSLSSAKGRCRAVLEPLPHRRRLAGVPGG 60	

Qy	61	EKSNNDYEDRBEQOLKAWMSMILPARPQSEYADTRYFODMABRLILPPKASVST	120
Db	61	EKCSNNNDYEDRBEQOLKAWMSMILPARPQSEYADTRYFODMABRLILPPKASVST	120
Qy	121	ERQTRDVMTOLEEDKPTFKFDVQSOFKGFKYKIKTGLPPRPAITLPPKKYQPLPPA	180
Db	121	ERQTRDVMTOLEEDKPTFKFDVQSOFKGFKYKIKTGLPPRPAITLPPKKYQPLPPA	180
Qy	181	PPRESSAYFARKPTFEYQVRQPRORSKADFSRYLGAEEESHNOTKRESSCPSSNONTOKS	240
Db	181	PPRESSAYFARKPTFEYQVRQPRORSKADFSRYLGAEEESHNOTKRESSCPSSNONTOKS	240
Qy	241	PPRAIASSSYMGKKSIAQRDHTGSMQNCPPAORCOAAASHSPRMILPYENTISEKPDPTKPD	300
Db	241	PPRAIASSSYMGKKSIAQRDHTGSMQNCPPAORCOAAASHSPRMILPYENTISEKPDPTKPD	300
Qy	301	EKDVMQNMWYIGEXSRQAVEDVLKEMKNDGFLVRDOSTSKAEPLYLVYFYCNKYNNK	360
Db	301	EKDVMQNMWYIGEXSRQAVEDVLKEMKNDGFLVRDOSTSKAEPLYLVYFYCNKYNNK	360
Qy	361	IRFLESNOQFALGTGLRGNEFMFDSVEDIIEHYTYFFPLILLIDGDKAARRKQCLTQPLPL	420
Db	361	IRFLESNOQFALGTGLRGNEFMFDSVEDIIEHYTYFFPLILLIDGDKAARRKQCLTQPLPL	420
Qy	421	ARLLLTQYSQALHE	435
Db	421	ARLLLTQYSQALHE	435

RESULT 2
Q9QZE2
ID Q9QZE2 PRELIMINARY; PRT; 435 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CLNK.
GN CLNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheta; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20029811; PubMed=10562326;
RA Cao M.Y., Davidson D., Yu J., Latour S., Velliette A.;
RT "Clnk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
RT Stimulated Hemopoietic Cells.";
RL J. Exp. Med. 190:1527-1534(1999).
DR EMBL: AF187819; AF14299.1; -
DR HSSP: P23727; P2NB.
DR MGI:1351468; CLNK.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PSS0001; SH2; 1.
SO SEQUENCE 435 AA; 49492 MW; 5CD27EC971FCC0A5 CRC64;

Query Match	99.5%	Score 2305	DB 11	Length 435
Best Local Similarly	99.58%	Pred. No. 6.5e-179		
Matches 433	Conservative 0	Mismatches 2	Indels 0	Gaps 0

[illegible]

Qy	121	ERQTRDVMTOLEEDYDKTFKFDVNSQSRFKGKYYIKTKTLP,PPRPRAIT,LPKKYQ,LPRA	160
Db	121	ERQTRDVMTHLEEDYDKTFKFDVNSQSRFKGKYYIKTKTLP,PPRPRAIT,LPKKYQ,LPRA	160
Qy	181	PPRESSAFAFKPTPEVQQRGRORSKADFSRYLGAEEESHQKTPRESSCPSSNQTKS	240
Db	181	PPRESSAFAFKPTPEVQQRGRORSKADFSRYLGAEEESHQKTPRESSCPSSNQTKS	240
Qy	241	PPRAISSSYMGKHSIQARDHTGSMQHCPAORCOAAASHSPRMLPYENTNSEKPPDDTKPD	300
Db	241	PPRAISSSYMGKHSIQARDHTGSMQHCPAORCOAAASHSPRMLPYENTNSEKPPDDTKPD	300
Qy	301	EKDVMQNMWTIGEXYSRQAVEDEVLMKENKDGTFVLVBCDSTKSKAEPYLVVYFGNKYYNVK	360
Db	301	EKDVMQNMWTIGEXYSRQAVEDEVLMKENKDGTFVLVBCDSTKSKAEPYLVVYFGNKYYNVK	360
Qy	361	IRFLESNQOFPALGTGLRGNEKMFDSVEDTIEHYTYFPLLLIDGKDKAARRKQCYLTQPL	420
Db	361	IRFLESNQOFPALGTGLRGNEKMFDSVEDTIEHYTYFPLLLIDGKDKAARRKQCYLTQPL	420
Qy	421	ARLLTOYSSQALHE	435
Db	421	ARLLTOYSSQALHE	435

RESULT 3
Q9P2U9
ID Q9P2U9 PRELIMINARY; PRT; 376 AA.

```

DE MIST.
GN MIST.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20208989; PubMed=10744659;
RA Goltisuka R., Kanazashi H., Sasanoana H., Fujiwara Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clink involved in Ige
RT receptor-mediated mast cell degranulation.";
RL Int. Immunol. 12:573-580(2000).
DR EMBL: AB032269; BAA96241.1; -.
DR HSSP; P23727; 2PNE.
DR InterPro: IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PRO0401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
FT NON_TER 1
FT TER 1
SQ SEQUENCE 376 AA; 43542 MW; 3492421B629B028B2 CRC64;

```

Query Match	48.98	Score 1132	DB 4	Length 376
Best Local Similarity	60.98	Pred. No. 8.2e-84		
Matches 229	Conservative 42	Mismatches 101	Indels 4	Gaps 3

```
Qy 19 F0N5L1KNSBSP1SS5KGCRAVL6PLPHRNLGVLGVGGEKSNNOYEDPEROLK 78
Dy 1 F0N5L1KNSBSP1NSTG0YORNNKPLDMEKNFAVLGDGANGSHDDYDDPELME60
Qy 79 AWP5KILPAP10E5EADRYEFQDM5EAPLLP5KASVSTERQTRDVMTOLEEDK6P 138
Dy 61 TW05KILPAP10E5EADHYFKVAMDPLPLDTYSISIQPTMNTQ-TLELRDKE 119
Qy 139 TFKDVR5QR5FGFKYTKINKTPLRPPRAITLPPKKYOLPDPAPBESSAYFAKP7FEY 198
```



```
Db 120 ISRDVRSQNIKGDASVKNKRIPLPPRLITLPPKTYQPLP-EPSSRPLLSQRTPEV 178
Qy 199 QGPRORSADKFSRYLGAEEESHQTKPESSCPSSNONQNSPPALASSYMPGHSIOA 258
Db 179 QGMPQSISLRLSEVLEAEKYPHNOQRKPESTHLENQTOEIPALISSSSFTSNHVSQN 238
Qy 259 RDHTSMQHCPAQRQAAASHSP--RMLPYENTSEKPDPTKPDKDVQWONWYIGESR 316
Db 239 RDHRGMOPCSPQRCQSPASCSPHENILPYKYSWPRPPKRDKDVQWONWYIGESR 298
Qy 317 QAVEVLKMKENKGTFLVRCSTKSKAPRYLVFYGKKNVYKIRFLESNOQFALGTGL 376
Db 299 QAVEAPFKENKDGSLVRCSTKSKKEERYLVAFYENKYNVYKIRFLERNQFALGTGL 358
Qy 377 RGENEMFVEDITIEHY 392
Db 359 RGDEKFDSEVEDITIEHY 374

RESULT 4
Q922M0 PRELIMINARY: PRT: 533 AA.
ID Q922M0;
AC Q922M0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lymphocyte cytosolic protein 2.
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC006948; AA06948.1;
DR MGD; MGI:1321402; LCP2.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 533 AA; 60238 MW; 50AEA025EF0DAD01 CRC64;

Query Match 17.7%; Score 410.5; DB 11; Length 533;
Best Local Similarity 29.7%; Pred. No. 4,4e-25;
Matches 148; Conservative 50; Mismatches 176; Indels 125; Gaps 20;

Qy 14 FGLRFQNSIL-----KNSWPSLSSAGRCRAVLEPLPDHRRNLG----- 56
Db 57 FPLRLPILSKLSQDINKNERSITRKQIRPLEFTESHEDDGGWSEFEDYESPN 116
Qy 57 --VPGSEKNSNDYEDP--EFOLL-----KAMPS-----MKILPARPIQ--ESEY 96
Db 117 DDDPDEG---DDGDYESPNEEQALVDADADYEPSPSNNEALQSSILPNSFHTNSMY 173
Qy 97 ADTRRYQDMMEAPLLP--PKASY----- 118
Db 174 IDRPYGGKYSQPPVPLRKPALDPLPTGRNHSPLSPPHNHEEPSRSNNKTAKIPAP 233
Qy 119 STERQTRDVMTQLLEVD-----KPTFKDYRSQRFKGYKTKIKTKPLPPRAPITL 170
Db 234 SIDRSRTPRLDRSLAPLRDPRFLGKKRPSPDSAPVAGREHLPIKIQPLP--PAMDR 290
Qy 171 PKTYQPLPAPPESSAYFAPKPTFFEVQGRORSADKSRVLGAEEESHQTKPESSC 230
Db 291 HERNERLGPVYTR-----KPPVRHGRGPPDRR-----NDDVDVHORLPPQSL 334
Qy 231 PSSNONQKSPAIASS-----SYMPGKHSIQARDHTGSMQHC--PAQCOAAASHSP- 281
Db 335 PSMSSNTFPSSRVSOPSSKNTFPLAHMPGARS---ESNIGQOQSSALPPTYSQGGNRPPL 391
Qy 282 ----RMLPYENTSEKPDPTKPDKDV--WQNEWYIGESYRQAVDVLMKENKDGTFVLRD 336
```

```
Db 392 RSEGNLPLPYPN--RRQPPSPGEEETPLDEEMVYSITRREAEALUKINQDGFELRD 449
Qy 337 CSTSKAPRYLVVYFGKKNVYKIRFLESNOQFALGTGLRGENMFSDVEDITIEHYTPP 396
Db 450 SSKTKANNPYVLMVLYKDKYVNIQIRYQESQVYLLGTGLGKEDFSLSDIIVFRKMP 509
Qy 397 ILLIDGKDKAARRKQCYLT 415
Db 510 LLLIDGKNRGS--RYOCTLT 527

RESULT 5
Q920L0 PRELIMINARY: PRT: 525 AA.
ID Q920L0;
AC Q920L0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE SLP-76 adaptor protein.
GN SLP76.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fujii Y., Goltzuka G., Wakahara S., Hara T., Nakao T.;
RT "Molecular cloning of SLP-76, a 76-kDa adaptor protein in RBL-2H3."
RL Submitted (OCR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB072980; BAB71779.1;
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 525 AA; 59521 MW; 0B1CDD40DABE7372 CRC64;

Query Match 17.7%; Score 410; DB 11; Length 525;
Best Local Similarity 31.9%; Pred. No. 4,8e-25;
Matches 127; Conservative 51; Mismatches 148; Indels 72; Gaps 16;

Qy 68 DYEDPEFOLLKAMPSSKILPAPRT--QSEIYADTRYQDMMEAP-----LLLP--- 114
Db 144 DYEPSPNDEEALQN--SILPAKFPPTNSMYIDRPTGKVSQPPVPPQRPMAALPPLPT 202
Qy 115 -----KASVTERQTRDVMTQLEVDKPTFEDVSQRRKGKTKYKINKTPL---- 161
Db 203 GRNHSNHEEPSRNHNKTKAPSIDNSTKPLD--NSLA-----PPLDREPLGKK 254
Qy 162 PPPRAPITLPPKTYQ--LPPAP-----EESAYFAPKPTFFEVQGRORSADK 210
Db 255 PSDKPSAPVAGREHLPIKQRPPLPAMDRIERNRIRPLTARKKPPVRHGRGPPDRD---- 310
Qy 211 SRVLGAEEESHQTKPESSCPSSNON-----TQSPRALASSYMPGKHSIQARDHTGS 264
Db 311 ----NDEDVDHORLPPHPSLPSMSNTFPSSRSPKPKNTFPLPHMGALS---ESNIG 363
Qy 265 MQHC--PAQCOAAASHSP-----RMLPYENTSEKPDPTKPDKDVQWONWYIGESRQ 317
Db 364 QQSASLPSITFSQSPSSRPVRYNRYNPLPVPNRPOP--PSPGEEESPLDEEMVYSITTR 422
Qy 318 AVEVLKMKENKDGTFVLRDCSTKSKAPRYLVVYFGKKNVYKIRFLESNOQFALGTGLR 377
Db 423 EAELALRKINQDGFVLRDSSKTKVNNPYVLMYKDKYVNIQIRYQESQVYLLGTGLR 482
Qy 378 GNEMFSDVEDITIEHYTPPILLIDGDKAARRKQCYLT 415
Db 483 GKEDFLSVSDIIVFRKMPILLIDGKNRGS--RYOCTLT 519

RESULT 6
Q9D607 PRELIMINARY: PRT: 530 AA.
ID Q9D607
```

AC 09D607;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE SLP-76 adaptor protein.
 CN SLP76.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxId=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=THYMUS;
 RX MEDLINE=20451093; PubMed=10993915;
 RA Ishii M., Kurosaki M., Inabe K., Chan A.C., Sugamura K., Kurosaki T.;
 RT "Involvement of LAT, Gads, and Grb2 in Compartmentation of SLP-76 to
 the Plasma Membrane."
 RL J. Exp. Med. 192:847-856(2000).
 DR EMBL: AF226988; AAC18493.1;
 DR HSSP: P12931; 1SHD.
 DR Interpro: IPR001660; SAM.
 DR Interpro: IPR000980; SH2.
 DR Pfam: PF00017; SH2.1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000093; SH2.1.
 DR SMART: SM00454; SAM; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2.1.
 SQ SEQUENCE 530 AA; 59580 MW; CF6CBCC8D79909A CRC64;

Query Match 16.5%; Score 381; DB 13; Length 530;
 Best Local Similarity 29.8%; Pred. No. 1.1e-22;

Matches 123; Conservative 54; Mismatches 151; Indels 85; Gaps 16;

QY 61 EKCSNND-YEDPEFOLKAMPKMLPAPPI-OESYATRTFQDMMEAPLLEPPASV 118
 DB 136 EEAHSDSDGTEPPPSNNDDEAHNV-IFPAKSLANNIDYIDRPTSRSSHQPPVPPQRP 194
 QY 119 STERQDVTMTQLEEVDPKFDVRSQRFKGYTKINKTLPAPPAITL----- 170
 DB 195 SPAPASFGGASL-----PAFPLPENNDNRNK-----PSKPPASIDRSTKPPDLR 242
 QY 171 -----PK-----YOLPPAPPESSAYFAKPTPE- 197
 DB 243 LGPPERENVPGRKPGHPEKLLTPOLRAGEQLAMMPKPPVPPSRYERGNPSPLKQI 302
 QY 198 -VQGPORSKADFSRVLGAEESH--HOTKPESSCPSSNQNT-----OKSPALIASSY 249
 DB 303 PVAGMAQOKRPE-----EEDHITPRAVPOISLPYSSNTFPKSKIKAPKPGNS 354
 QY 250 MEGKHSIQANDHTGSMOHCFAORCAQAASHSPM-----LPYENTNSEKPDPTKPD 301
 DB 355 IPGAESARSLASAGSL---PPRPPLGNSSPSRGTAADLRPLRPIPSRQAHQNTNEEDE 411
 QY 302 KDWONMWYGEYSRQAEVDYLMKENKDGTFILVDCSTSKAPYLVVYFGKVVNWKI 361
 DB 412 -DLENDEMYAVYSRPAEALAKINKDGFELVDSRSKTTTPYLVMLYKXKVINQI 470
 QY 362 RFLESNOQFALGTGLNGENMFDSVEDIIEHYTFPIILLIDGKKAARKQCYL 414
 DB 471 RYQEOQTYLTGLTGKKEDEFSVAHIIDYFOTPIILLIDGKRGSR-RNQCVL 522

RESULT 7
 08WV28 PRELIMINARY; PRT; 456 AA.
 AC 08WV28;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE B-cell linker.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TONSIL;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC018906; AAH18906.1;
 DR Interpro: IPR000980; SH2.
 DR Pfam: PF00017; SH2.1.
 DR PRODOM: PD000093; SH2.1.
 DR SMART: SM00252; SH2.1.
 DR PROSITE: PS50001; SH2.1.
 SQ SEQUENCE 456 AA; 50465 MW; ED6D42A035D1792 CRC64;

Query Match 13.8%; Score 319; DB 4; Length 456;
 Best Local Similarity 25.0%; Pred. No. 9.7e-18;

Matches 99; Conservative 58; Mismatches 139; Indels 100; Gaps 12;

QY 64 NSNNYEDPEFOLKAMPKMLPAPPIQSEYADTRYFQDMMEAP--LLPPKASVST 120
 DB 91 NADDSYEPPEVE---QETRPVHPALPFARGEYIDNRSSQ--RHSPFSKTLPSKPSMPS 144
 QY 121 ERQTRQVMTQLEEVDPKFDVRSQRFKGYTKINKTLPAPPAITLPPKYYQPLPA 180
 DB 145 EKARLTSTLPAALOKPY-----PPKRGLEDDADYVVP 182
 QY 181 PRESSAYFAKPTPEVQGRORSKADFSRVLGAEESHQHTKPESSCPSSNQNT--- 237
 DB 183 EDNDENYIHTESSSPPEKAPM-----VNRSTKNSSPAPSPETAGS 226
 QY 238 -----OKSPALIASSYMP--GKH-----STQARDHTSM---QHCFAORCAQA- 276
 DB 227 RNSGAMETSSPP-APSPPLPRAKKPTTPTTPVASQOMASSVCEKPTPAERHRGSS 285
 QY 277 ---ASHSPMLPYENTNSEKPD-----TKPDKDVQNE 308
 DB 286 HREQAVQSPVPPPAQOHQKPIPLRFTGEGNPTVDGLPSTSSSTISDEAGVLCRP 345
 QY 309 WYIGEYSRQAEVDYLMKENKDGTFILVDCSTSKAPYLVVYFGKVVNWKIRFLESNQ 368
 DB 346 WTAGACDRKSAEALHRSNMGDFELIRKSSGHSKQPYLVVFNKRVYNIPVRFTEATK 405
 QY 369 OFALGTGLNGENMFDSVEDIIEHYTFPIILLIDGKD 404
 DB 406 QYALGKRKNGEEFGSVAEIIRHNSPLVLDSON 441

RESULT 8
 075498 PRELIMINARY; PRT; 456 AA.
 AC 075498;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE B cell linker protein BLNK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98001722; PubMed=9341187;
 RA Fu C., Chan A.C.;
 RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
 interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
 receptor activation."
 RL J. Biol. Chem. 272:27362-27368(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98361304; PubMed=9697839;
 RA Fu C., Turck C.W., Kurosaki T., Chan A.C.;


```

RESULT 10
075499 PRELIMINARY; PRT; 433 AA.
ID 075499
AC 075499;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE B cell linker protein BLNK-S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation."
RL J. Biol. Chem. 272:27362-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turk C.W., Kurosaki T., Chan A.C.;
RT "BLNK: a central linker protein in B cell activation."
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20050956; PubMed=10583958;
RA Mlegishi Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
RA Campana D., Chan A.C., Conley M.E.;
RT "An essential role for BLNK in human B cell development."
RL Science 286:1954-1957(1999).
DR EMBL: AF068181; AAC39937.1;
DR EMBL: AF180756; AAF20383.1;
DR EMBL: AF180740; AAF20383.1; JOINED.
DR EMBL: AF180741; AAF20383.1; JOINED.
DR EMBL: AF180742; AAF20383.1; JOINED.
DR EMBL: AF180743; AAF20383.1; JOINED.
DR EMBL: AF180744; AAF20383.1; JOINED.
DR EMBL: AF180745; AAF20383.1; JOINED.
DR EMBL: AF180746; AAF20383.1; JOINED.
DR EMBL: AF180748; AAF20383.1; JOINED.
DR EMBL: AF180749; AAF20383.1; JOINED.
DR EMBL: AF180750; AAF20383.1; JOINED.
DR EMBL: AF180751; AAF20383.1; JOINED.
DR EMBL: AF180752; AAF20383.1; JOINED.
DR EMBL: AF180753; AAF20383.1; JOINED.
DR EMBL: AF180754; AAF20383.1; JOINED.
DR EMBL: AF180755; AAF20383.1; JOINED.
DR HSSP: P29354; 1BMB.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR ProDom: PD000093; SH2.1.
DR SMART: SM00252; SH2.1.
DR PROSITE: PS50001; SH2.1.
SQ SEQUENCE 433 AA; 48229 MW; 0B36FE9FCF5DC7DC CRC64;

Query Match 13.48; Score 309.5; DB 4; Length 433;
Best Local Similarity 25.38; Pred. No. 5.3e-17;
Matches 99; Conservative 51; Mismatches 127; Indels 115; Gaps 12;

```

```

DB 168 PKGLEDEADYVP-----VEDDENYIHPTESSPEPEGRNS 206
QY 237 ----TQSPPAIASSSYMBGKH-----SIQARDHTGSM---OHCPAORCOAA----- 276
DB 207 GAMETKSPPPAPPLPRGAKKPTPLKTTTPVAAQQNASSVCEKEIPAEHRGSSHQE 266
QY 277 ASHSPMLLYENTNSKPP-----IKPDEKDWQNEWTG 312
DB 267 AVQSPVPPPAOKIHKPIPLPFTFEGNPTVDGLPFSFSSNSTISEQAGVLCRPMVAG 326
QY 313 EYSROAVEDYLNKNGDTFLVDCSTKSKAPRYLVFYGKNVKNVIFLESNOQFAL 372
DB 327 ACDRKSAEALIRSNKDSGFLIRKSGHSGKQPTLYVFENKRVINIPVFIEATKQVAL 386
QY 373 GTGLNGENKMFSDVEDIIEHYTFPILLIDGKD 404
DB 387 GRKKNGEEYFGSVAETIRHQHSPVLIDSON 418

RESULT 11
090UN3 PRELIMINARY; PRT; 457 AA.
ID 090UN3
AC 090UN3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE SH2-containing leukocyte protein 65 (Lymphocyte antigen 57).
GN SLP-65 OR BASH OR LY57.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/C; TISSUE-LYMPHOID;
RX MEDLINE=98372771; PubMed=9705962;
RA Wienands J., Schweikert J., Wollschlaed B., Junna H., Nielsen P.J.,
RA Reth M.;
RT "SLP-65: A new signalling component in B lymphocytes which requires
RT expression of the antigen receptor for phosphorylation."
RL J. Exp. Med. 188:791-795(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/C; TISSUE-LYMPHOID;
RA Wienands J., Labollette O., Reth M.;
RT "Evidence for a preformed transducer complex organized by the B cell
RT antigen receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 938:7865-7870(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX Okamoto N., Hayashi K., Tsuji S., Gotsuka R., Kitamura D.;
RT "BASH: B lymphocyte adaptor protein containing SH2 domain."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX TRANSDON-ETN;
RA Nielsen P.J., Guenet J.L.;
RT "The murine SLP-65 gene."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y17159; CAA7666.1;
DR EMBL: AB015290; BAA34944.1;
DR EMBL: AJ298054; CAC18565.1;
DR HSSP: P23727; 1BFI.
DR MCD: MGI:96878; LY57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2.1.
DR SMART: SM00252; SH2.1.
DR PROSITE: PS50001; SH2.1.
KW B-cell.
FT CONFLICT 133 133 S -> N (IN REF. 2).

```



```
Db 201 SLDEPWTE-BEGD-----GSDHPYNSIPSKMPPPGCFIDTRLKPRPHAP-- 244
QY 184 ESSAYEAPK-----PTPE-VQGRPRORSAKDFSRVLGAEEESHQTKPESSCP 231
Db 245 -DTAOFAGKEQTYQGRHLDGTFEGEDWQOTPLROGSSD---IYSTPEGKLVHAPTEGEAP 299
QY 232 SSNNTQKSPPALASSSYMFGKHSIQARDHTGSMQHCPAQRCAASHSPR-----MLPY 286
Db 300 -TYVNTQOIPR-----QANPA-AVSSAESSPRKDLFDMPF 333
QY 287 ENTNSEKP-----DPTKPDEK---DWMQNEWYIGEYSROAVEDVLMKE 326
Db 334 EDALKNQPLGPVLSKASVCEISPVSPRAPDAKMLEELQAFETWYQGENSKREAEGLL--- 390
QY 327 NKDGTFLVRDCSTKSKAEP--YVLVIFYGNKVYVKIRFLESNOQFALGTGLRGENEMFDS 384
Db 391 EKDDGDFLVR---KSTNPGSFVLTGMHNGQAKHLLVDP-----GTERTKDRVFD 439
QY 385 VEDIIEHY--TYPIILLIDGKDKAARRKOCYLOPL 418
Db 440 ISHLINHLESPLIV-----SAGSELC-LOOPV 467
```

Search completed: April 21, 2003, 12:37:55
Job time : 33.6461 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:32:18, Search time 31.6461 Seconds

(without alignments)
2832.275 Million cell updates/sec

Title: US-09-856-061-2

Perfect score: 2316

Sequence: 1 MTSQGNKRTTKEGFDLRFQ.....QPLPLALLTQYSSQALHE 435

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database:

SPREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	11 Q9JMJ3	Q9JMJ3 mus musculus
2	2305	99.5	435	11 Q9QZ22	Q9QZ22 mus musculus
3	1132	48.9	376	4 Q9P209	Q9P209 homo sapien
4	410.5	17.7	533	11 Q922M0	Q922M0 mus musculus
5	410	17.7	525	11 Q920L0	Q920L0 rattus norv
6	381	16.5	530	13 Q9DG07	Q9DG07 gallus gall
7	319	13.8	456	4 Q8WY28	Q8WY28 homo sapien
8	319	13.8	456	4 Q7S498	Q7S498 homo sapien
9	315.5	13.6	552	13 Q9YGC1	Q9YGC1 gallus gall
10	309.5	13.4	433	4 Q75499	Q75499 homo sapien
11	297.5	12.8	457	11 Q9QUN3	Q9QUN3 mus musculus
12	287.5	12.4	457	11 Q88504	Q88504 mus musculus
13	219.5	9.5	297	11 Q9D413	Q9D413 mus musculus
14	188.5	8.1	516	11 Q54737	Q54737 mus musculus
15	167.5	7.2	471	4 Q9DCX5	Q9DCX5 homo sapien
16	165.5	7.1	594	4 Q92529	Q92529 homo sapien

17	165.5	7.1	594	4 Q8TAP2	Q8TAP2 homo sapien
18	150.5	6.5	486	4 Q9UCX4	Q9UCX4 homo sapien
19	150.5	6.5	812	15 Q85466	Q85466 y73 sarcoma
20	143.5	6.2	662	5 Q9N355	Q9N355 caenorhabdi
21	141.5	6.1	474	11 Q61120	Q61120 mus musculus
22	141	6.1	728	4 Q9UPH9	Q9UPH9 homo sapien
23	139.5	6.0	286	5 Q95PX0	Q95PX0 caenorhabdi
24	139.5	6.0	398	5 Q95PW9	Q95PW9 caenorhabdi
25	138.5	6.0	335	4 Q96CZ8	Q96CZ8 homo sapien
26	138.5	6.0	926	4 Q9H0K1	Q9H0K1 homo sapien
27	137.5	5.9	594	11 Q70143	Q70143 rattus norv
28	136.5	5.9	477	11 Q70142	Q70142 rattus norv
29	134	5.8	309	11 Q9QVW0	Q9QVW0 mus sp. shb
30	134	5.8	461	11 Q63789	Q63789 rattus norv
31	134	5.8	1599	5 Q18892	Q18892 caenorhabdi
32	133.5	5.8	217	11 Q9CX99	Q9CX99 mus musculus
33	132.5	5.7	904	5 Q9VMJ4	Q9VMJ4 drosophila
34	132.5	5.7	944	5 Q8R9B3	Q8R9B3 drosophila
35	130.5	5.6	722	13 Q8UUY2	Q8UUY2 xenopus lae
36	129	5.6	559	11 Q91Z52	Q91Z52 mus musculus
37	129	5.6	645	6 Q95U15	Q95U15 macaca fasc
38	128.5	5.5	112	11 P70305	P70305 mus musculus
39	128.5	5.5	393	2 Q93CA3	Q93CA3 treponema p
40	128.5	5.5	854	11 Q9QZS8	Q9QZS8 mus musculus
41	128.5	5.5	1430	11 Q8VHK2	Q8VHK2 rattus norv
42	128	5.5	623	6 Q95J17	Q95J17 macaca fasc
43	128	5.5	651	11 P97504	P97504 mus musculus
44	128	5.5	1160	11 Q9ESC8	Q9ESC8 mus musculus
45	128	5.5	4498	13 Q93291	Q93291 tugu rubrip

ALIGNMENTS

RESULT 1	Q9JMJ3	PRELIMINARY:	PRT:	435 AA.
AC	Q9JMJ3			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	MIST.			
OS	CLNK OR MIST.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-20208989; PubMed-10744659;			
RA	Goitsuka R., Kanazashi H., Sasanna H., Fujimura Y., Hida Y.,			
RA	Tatsuno A., Ra C., Hayashi K., Kitamura D.,			
RT	"A BASH/SLP-76-related adaptor protein MIST/Clnk involved in Ige			
RT	receptor-mediated mast cell degranulation.";			
RL	Int. Immunol. 12:573-580(2000).			
DR	EMBL: AB021220; BAA96240.1; -			
DR	HSSP: P23727; 2PNB.			
DR	MGD; MGI:1351468; Clnk.			
DR	InterPro; IPR000980; SH2.			
DR	Pfam; PF00017; SH2; 1.			
DR	PRINTS; PR00401; SH2DOMAIN.			
DR	ProDom; PD000093; SH2; 1.			
DR	SMART; SM00252; SH2; 1.			
DR	PROSITE; PS50001; SH2; 1.			
SQ	SEQUENCE 435 AA; 49513 MW; 4D1BD3E2F0C61ED6 CRC64;			
Query Match	100.0%; Score 2316; DB 11; Length 435;			
Best local similarity	100.0%; Pred. No. 8.3e-180;			
Matches 435; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MTSQGNKRTTKEGFDLRFQVNSLKNRSPSLSSAKKRCRAVLEPLPDHRRNLACVPGG 60			
DB	1 MTSQGNKRTTKEGFDLRFQVNSLKNRSPSLSSAKKRCRAVLEPLPDHRRNLACVPGG 60			


```

OY 61 EKCSNNNDYEDPEFOLKAMPMSKILPARPIQSEYADTRFQDMMEAPLLPPKASVST 120
    |||||
DB 61 EKCSNNNDYEDPEFOLKAMPMSKILPARPIQSEYADTRFQDMMEAPLLPPKASVST 120
OY 121 ERQTRDVRMTQLEEVDPKPTFKDVRSGRFGKFKYTKINKTLPPEPPRAITLPPKKYQPLPPA 180
    |||||
DB 121 ERQTRDVRMTQLEEVDPKPTFKDVRSGRFGKFKYTKINKTLPPEPPRAITLPPKKYQPLPPA 180
OY 181 PPESSAYFAKPTPEVQGRPRQSAKDFSRVLGAEEESHQTKPESSCPSSNQNTQKS 240
    |||||
DB 181 PPESSAYFAKPTPEVQGRPRQSAKDFSRVLGAEEESHQTKPESSCPSSNQNTQKS 240
OY 241 PPAIASSYMPGKHSIQADHTGSMQCHPQRCQAASHSPMLPYENTNSEKPDPTKPD 300
    |||||
DB 241 PPAIASSYMPGKHSIQADHTGSMQCHPQRCQAASHSPMLPYENTNSEKPDPTKPD 300
OY 301 EKDVQWQNEWYIGEYSRQAVEDVLMEKNKDGTFVLRDCSTKSAEPYLVVYFGNKYVYVY 360
    |||||
DB 301 EKDVQWQNEWYIGEYSRQAVEDVLMEKNKDGTFVLRDCSTKSAEPYLVVYFGNKYVYVY 360
OY 361 IRFLESNQOPFALGTGLRGNMEDSVEDIIEHTYTFPILLIDGKDKAARRKQCYLTQPLPL 420
    |||||
DB 361 IRFLESNQOPFALGTGLRGNMEDSVEDIIEHTYTFPILLIDGKDKAARRKQCYLTQPLPL 420
OY 421 ARLLLTQYSSQALHE 435
    |||||
DB 421 ARLLLTQYSSQALHE 435

```

RESULT 2

```

OY 09QZ2 PRELIMINARY: PRT: 435 AA.
AC 09QZ2:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CLNK.
GN CLNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20029811; PubMed=10562326;
RA Cao M.Y., Davidson D., Yu J., Latour S., Veillette A.;
RT "Clunk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytochrome-
RL J. Exp. Med. 190:1527-1534(1999).
DR EMBL; AF187819; AAF14299.1;
DR HSSP; P23727; 2PMB.
DR MGD; MGI:1351468; Clnk.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
SQ SEQUENCE 435 AA; 49492 MW; 5CD27EC971FC0EAS CRC64;

```

Query Match 99.5%; Score 2305; DB 11; Length 435;
 Best Local Similarity 99.5%; Pred. No. 6.5e-179;
 Matches 433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 MTSQGNKRTTEGEGDLRFQVNSILKNSWPSLSASAKRCRAVLEPLPDHRRNLAVGPGG 60
    |||||
DB 1 MTSQGNKRTTEGEGDLRFQVNSILKNSWPSLSASAKRCRAVLEPLPDHRRNLAVGPGG 60
OY 61 EKCSNNNDYEDPEFOLKAMPMSKILPARPIQSEYADTRFQDMMEAPLLPPKASVST 120
    |||||
DB 61 EKCSNNNDYEDPEFOLKAMPMSKILPARPIQSEYADTRFQDMMEAPLLPPKASVST 120

```

```

OY 121 ERQTRDVRMTQLEEVDPKPTFKDVRSGRFGKFKYTKINKTLPPEPPRAITLPPKKYQPLPPA 180
    |||||
DB 121 ERQTRDVRMTQLEEVDPKPTFKDVRSGRFGKFKYTKINKTLPPEPPRAITLPPKKYQPLPPA 180
OY 181 PPESSAYFAKPTPEVQGRPRQSAKDFSRVLGAEEESHQTKPESSCPSSNQNTQKS 24
    |||||
DB 181 PPESSAYFAKPTPEVQGRPRQSAKDFSRVLGAEEESHQTKPESSCPSSNQNTQKS 2
OY 241 PPAIASSYMPGKHSIQADHTGSMQCHPQRCQAASHSPMLPYENTNSEKPDPTKPD 300
    |||||
DB 241 PPAIASSYMPGKHSIQADHTGSMQCHPQRCQAASHSPMLPYENTNSEKPDPTKPD 300
OY 301 EKDVQWQNEWYIGEYSRQAVEDVLMEKNKDGTFVLRDCSTKSAEPYLVVYFGNKYVYVY 360
    |||||
DB 301 EKDVQWQNEWYIGEYSRQAVEDVLMEKNKDGTFVLRDCSTKSAEPYLVVYFGNKYVYVY 360
OY 361 IRFLESNQOPFALGTGLRGNMEDSVEDIIEHTYTFPILLIDGKDKAARRKQCYLTQPLPL 420
    |||||
DB 361 IRFLESNQOPFALGTGLRGNMEDSVEDIIEHTYTFPILLIDGKDKAARRKQCYLTQPLPL 420
OY 421 ARLLLTQYSSQALHE 435
    |||||
DB 421 ARLLLTQYSSQALHE 435

```

RESULT 3

```

OY 09P209 PRELIMINARY: PRT: 376 AA.
AC 09P209:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE M1SF (Fragment).
GN M1SF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20208989; PubMed=10744659;
RA Goitsuka R., Kanazashi H., Sasamura H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein M1SF/Clunk involved in Ige
RL Int. Immunol. 12:573-580(2000).
DR EMBL; AB032369; BAA96241.1;
DR HSSP; P23727; 2PMB.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
FT NON_TER 1
FT NON_TER 376
SQ SEQUENCE 376 AA; 43542 MW; 389421B629B028B2 CRC64;

```

Query Match 48.9%; Score 1132; DB 4; Length 376;
 Best Local Similarity 60.9%; Pred. No. 8.2e-64;
 Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

```

OY 19 FQVNSILKNSWPSLSASAKRCRAVLEPLPDHRRNLAVGPGGKNSNNNDYEDPEFOLK 78
    |||||
DB 1 FQVNSILKNSWPSLSASAKRCRAVLEPLPDHRRNLAVGPGGKNSNNNDYEDPEFOLK 78
OY 79 AWPMSKILPARPIQSEYADTRFQDMMEAPLLPPKASVSTQTRDVRMTQLEEVDPK 138
    |||||
DB 79 AWPMSKILPARPIQSEYADTRFQDMMEAPLLPPKASVSTQTRDVRMTQLEEVDPK 138
OY 139 TFKDVNSQKRGKRYKINKTLPPEPPRAITLPPKKYQPLPPAPPESSAYFAKPTPEV 198
    |||||
DB 139 TFKDVNSQKRGKRYKINKTLPPEPPRAITLPPKKYQPLPPAPPESSAYFAKPTPEV 198

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:32:18 ; Search time 27.3539 Seconds

(without alignments)
2832.275 Million cell updates/sec

Title: US-09-856-061-4

Perfect score: 2020

Sequence: 1 FQNSLPKRNKSPRINSATG.....GLNGDEKPDSPVEDIEHYKN 376

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	2020	100.0	376	4 Q9P2U9	Q9P2U9 homo sapien
2	1132	56.0	435	11 Q9JMJ3	Q9JMJ3 mus musculi
3	1132	56.0	435	11 Q9QZL2	Q9QZL2 mus musculi
4	334	16.5	552	13 Q9YGC1	Q9YGC1 gallus gall
5	317.5	15.7	530	13 Q9DGO7	Q9DGO7 gallus gall
6	316.5	15.7	525	11 Q92DLO	Q92DLO rattus norv
7	314.5	15.6	533	11 Q922M0	Q922M0 mus musculi
8	289.5	14.3	433	4 Q75499	Q75499 homo sapien
9	289	14.3	456	4 Q8WV28	Q8WV28 homo sapien
10	282	14.0	457	11 Q9QUN3	Q9QUN3 mus musculi
11	282	14.0	457	11 Q88504	Q88504 mus musculi
12	273	13.5	457	11 Q9D413	Q9D413 mus musculi
13	177	8.8	297	11 Q54737	Q54737 mus musculi
14	153	7.6	516	11 P70304	P70304 mus musculi
15	147	7.3	724	11 P70304	P70304 mus musculi
16	144.5	7.2	1616	4 O15054	O15054 homo sapien

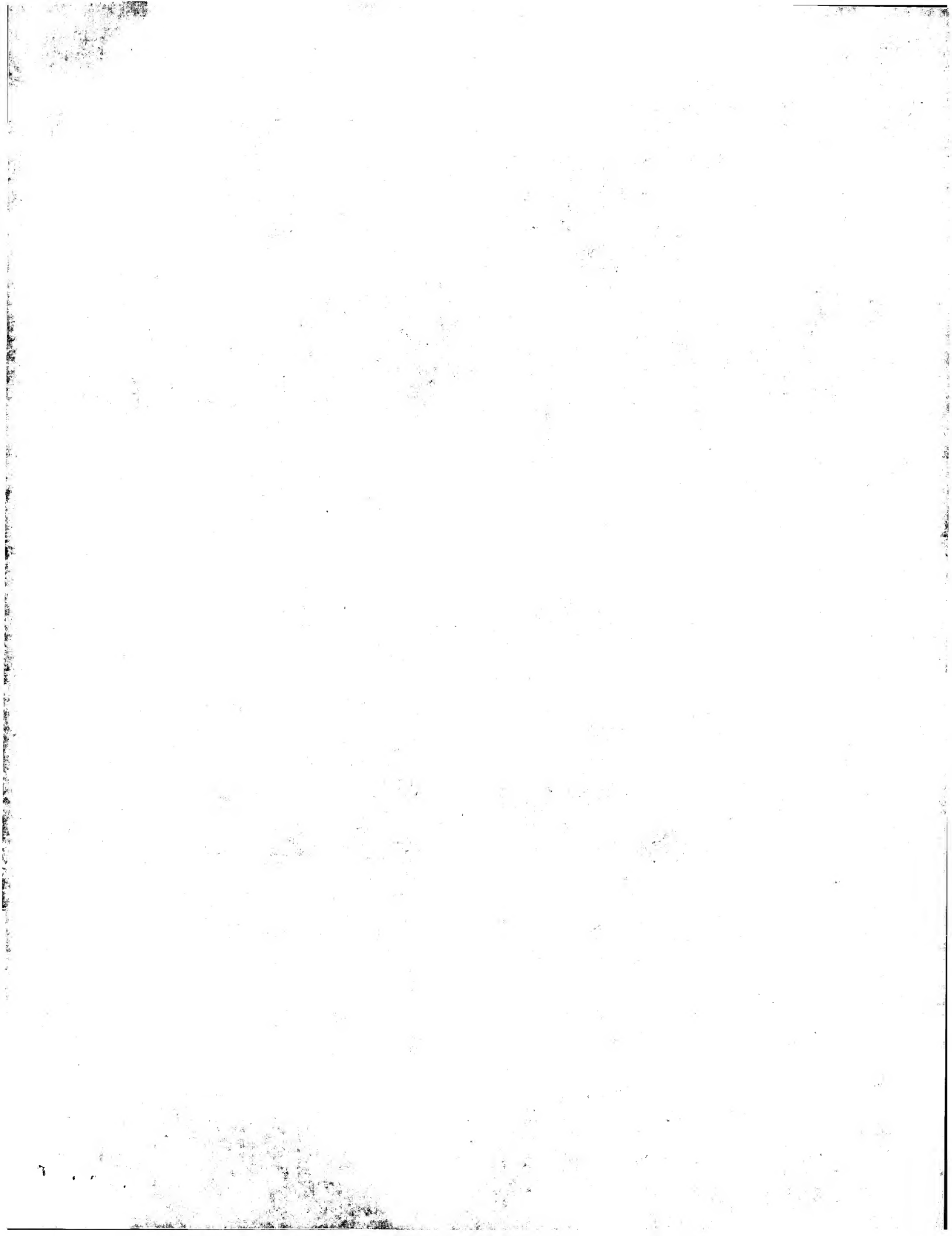
17	142	7.0	722	13 Q8UUN2	Q8UUN2 xenopus lae
18	138.5	6.9	981	15 Q92809	Q92809 abelson mur.
19	135	6.7	559	11 Q91252	Q91252 mus musculi
20	134.5	6.7	662	5 Q9N355	Q9N355 caenorhabdi
21	132	6.5	112	11 P70305	P70305 mus musculi
22	132	6.5	764	5 Q24708	Q24708 drosophila
23	131.5	6.5	471	4 Q9UCX5	Q9UCX5 homo sapien
24	131.5	6.5	485	3 Q9UTP6	Q9UTP6 schizosacch
25	130.5	6.5	548	5 Q9VPU1	Q9VPU1 drosophila
26	129.5	6.4	594	4 Q92529	Q92529 homo sapien
27	129.5	6.4	594	4 Q8RAP2	Q8RAP2 homo sapien
28	128	6.3	17352	5 Q95YM2	Q95YM2 procamburus
29	126.5	6.3	728	4 Q9UPH9	Q9UPH9 homo sapien
30	126	6.2	377	11 Q55032	Q55032 mus musculi
31	126	6.2	377	13 P79956	P79956 xenopus lae
32	126	6.2	461	11 Q63789	Q63789 rattus norv
33	125.5	6.2	533	11 Q8VEN5	Q8VEN5 mus musculi
34	125.5	6.2	1258	11 Q8R457	Q8R457 mus musculi
35	125	6.2	269	11 Q8CT18	Q8CT18 mus musculi
36	125	6.2	377	11 Q92279	Q92279 mus musculi
37	125	6.2	377	11 Q99M51	Q99M51 mus musculi
38	125	6.2	460	4 Q96HD7	Q96HD7 homo sapien
39	125	6.2	595	13 Q92124	Q92124 xenopus lae
40	125	6.2	597	11 Q64509	Q64509 mus musculi
41	125	6.2	823	6 Q9TTY2	Q9TTY2 canis famli
42	124.5	6.2	217	11 Q9CX99	Q9CX99 mus musculi
43	124	6.1	410	5 Q24218	Q24218 drosophila
44	123.5	6.1	863	4 Q9UPV4	Q9UPV4 homo sapien
45	123	6.1	927	4 Q9UEG4	Q9UEG4 homo sapien

ALIGNMENTS

RESULT 1
Q9P2U9 PRELIMINARY: PRT: 376 AA.

AC Q9P2U9; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MST (Fragment).
GN MST.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20208989; PubMed=10744659;
RA Gotsuka R., Kanazashi H., Sasamura H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.,
RT "A BASH/SLP-76-related adaptor protein MST/Clnk involved in Ige
RT receptor-mediated mast cell degranulation.";
RL Int. Immunol. 12:573-580(2000).
DR EMBL: AB032369; BAA96241.1; -;
DR HSSP: P23727; 2PNB.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PSS0001; SH2; 1.
FT NON_TER 1
FT TER 1
SQ SEQUENCE 376 AA; 43542 MW; 389421B629B028B2 CRC64;

Query Match 100.0%; Score 2020; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.6e-158;
Matches 376; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 FQNSLPKRNKSPRINSATGQYQRMKPLDWERNFVAVIDGAKGSHSDYDDPELMEF 60
|||||



```

Db 1 FONFSLPKNRSPRINSATGOYORANKPLLDWERNFAVLGAKGSHDDYDDPELMEE 60
QY 61 TWOSIKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTRLEVDKPI 120
Db 61 TWOSIKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTRLEVDKPI 120
QY 121 SMDVASONIKGDASVAKKNIPLPPRPPLITLPPKQYPLPEPESRRPPLSGRHTPEVOG 180
Db 121 SMDVASONIKGDASVAKKNIPLPPRPPLITLPPKQYPLPEPESRRPPLSGRHTPEVOG 180
QY 181 MFSQSLNDLSEVLAEKVPHNOKPESTHLENTQOEIPLAIISSSFTTSHNSVQND 240
Db 181 MFSQSLNDLSEVLAEKVPHNOKPESTHLENTQOEIPLAIISSSFTTSHNSVQND 240
QY 241 HRGMOQPCSPORCOPPASCSPHENILPYKYSWRRPPFPKRSRDKVDQHNEMVIGEYSROA 300
Db 241 HRGMOQPCSPORCOPPASCSPHENILPYKYSWRRPPFPKRSRDKVDQHNEMVIGEYSROA 300
QY 301 VEEAFMKENKDGSLVRCSTKSKEEPLYLAVFENKYNVAKIRFLERNOQFALGTGLG 360
Db 301 VEEAFMKENKDGSLVRCSTKSKEEPLYLAVFENKYNVAKIRFLERNOQFALGTGLG 360
QY 361 DEKEDSVEDIIEHYKN 376
Db 361 DEKEDSVEDIIEHYKN 376

```

RESULT 2

```

Q9JMJ3 PRELIMINARY; PRT; 435 AA.
ID 09JMJ3
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE MIST.
GN CLNK OR MIST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=20208989; PubMed=10744655;
RA Gotsuka R., Kanazashi H., Sasakuma H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clunk involved in Ige
RT receptor-mediated mast cell degranulation.";
RL Int. Immunol. 12:573-580(2000).
DR EMBL: AB021220; BAA96240.1; -.
DR HSSP: P23727; 2PNB.
DR MGD: MGI:1351468; Clnk.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SSQDNCE 435 AA; 49513 MW; 4D1BD3E2FC061ED6 CRC64;

```

```

Query Match 56.0%; Score 1132; DB 11; Length 435;
Best Local Similarity 60.9%; Pred. No. 3.2e-85;
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

```

```

QY 1 FONFSLPKNRSPRINSATGOYORANKPLLDWERNFAVLGAKGSHDDYDDPELMEE 60
Db 19 FONVSLKNSWPSLSAKGRCAVLEPLDHRNRNLGAVGGEKCSNNNDYEDPEFQLK 78
QY 61 TWOSIKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTRLEVDKPI 119
Db 79 AWMSPKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTRLEVDKPI 138
QY 120 ISMDVASONIKGDASVAKKNIPLPPRPPLITLPPKQYPLPEPESRRPPLSGRHTPEV 178

```

```

Db 139 TFKDVRSQRFKFKYTKINKTLPPLPPPAITLPPKQYPLPEPESRRPPLSGRHTPEV 198
QY 179 QGMFSQSLNDLSEVLAEKVPHNOKPESTHLENTQOEIPLAIISSSFTTSHNSVQND 238
Db 199 QRGPRQSAKDPFSLVGAEEESHQOTPESSCPSSNONTQKSPALAISSSYMPGKHSIOA 258
QY 239 RDHFGMOQPCSPORCOPPASCSPHENILPYKYSWRRPPFPKRSRDKVDQHNEMVIGEYSR 298
Db 259 RDHFGMOQPCSPORCOPPASCSPHENILPYKYSWRRPPFPKRSRDKVDQHNEMVIGEYSR 298
QY 299 QAVEAFMKENKDGSLVRCSTKSKEEPLYLAVFENKYNVAKIRFLERNOQFALGTGL 358
Db 317 QAVEDVLKMKDKGTPLPDCSTKSKEEPLYLAVFENKYNVAKIRFLERNOQFALGTGL 376
QY 359 RGDEKEDSVEDIIEHY 374
Db 377 RGNEMFDSVEDIIEHY 392

```

RESULT 3

```

Q9QZE2 PRELIMINARY; PRT; 435 AA.
ID 09QZE2
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CLNK.
GN CLNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=20029811; PubMed=10562326;
RA Cao M.Y., Davidson D., Yu J., Latour S., Veillette A.;
RT "Clunk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
RT stimulated Hemopoietic Cells.";
RL J. Exp. Med. 190:1527-1534(1999).
DR EMBL: AF187819; AAF14299.1; -.
DR HSSP: P23727; 2PNB.
DR MGD: MGI:1351468; Clnk.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SSQDNCE 435 AA; 49492 MW; 5CD27EC971FCD6A5 CRC64;

```

```

Query Match 56.0%; Score 1132; DB 11; Length 435;
Best Local Similarity 60.9%; Pred. No. 3.2e-85;
Matches 229; Conservative 41; Mismatches 102; Indels 4; Gaps 3;

```

```

QY 1 FONFSLPKNRSPRINSATGOYORANKPLLDWERNFAVLGAKGSHDDYDDPELMEE 60
Db 19 FONVSLKNSWPSLSAKGRCAVLEPLDHRNRNLGAVGGEKCSNNNDYEDPEFQLK 78
QY 61 TWOSIKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTRLEVDKPI 119
Db 79 AWMSPKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTRLEVDKPI 138
QY 120 ISMDVASONIKGDASVAKKNIPLPPRPPLITLPPKQYPLPEPESRRPPLSGRHTPEV 178
Db 139 TFKDVRSQRFKFKYTKINKTLPPLPPPAITLPPKQYPLPEPESRRPPLSGRHTPEV 198
QY 179 QGMFSQSLNDLSEVLAEKVPHNOKPESTHLENTQOEIPLAIISSSFTTSHNSVQND 238
Db 199 QRGPRQSAKDPFSLVGAEEESHQOTPESSCPSSNONTQKSPALAISSSYMPGKHSIOA 258
QY 239 RDHFGMOQPCSPORCOPPASCSPHENILPYKYSWRRPPFPKRSRDKVDQHNEMVIGEYSR 298

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:32:18 ; Search time 27.3539 Seconds
(without alignments)
2832.275 Million cell updates/sec

Title: US-09-856-061-4
Perfect score: 2020
Sequence: 1 FQNFSLPKRNSWPRINSATG.....GLNGDEKFDSDIEDIEHYKN 376

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SPTREMBL_21.*

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_ricent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2020	100.0	376	4	Q9P2U9
2	1132	56.0	435	11	Q9JMJ3
3	1132	56.0	435	11	Q9QZC2
4	334	16.5	552	13	Q9YGC1
5	317.5	15.7	530	13	Q9DG07
6	316.5	15.7	525	11	Q920L0
7	314.5	15.6	533	11	Q922M0
8	289.5	14.3	433	4	Q75499
9	289	14.3	436	4	Q8WV28
10	289	14.3	456	4	Q75498
11	282	14.0	457	11	Q9QUN3
12	273	13.5	457	11	Q88504
13	177	8.8	297	11	Q9D413
14	153	7.6	516	11	Q54737
15	147	7.3	724	11	P70304
16	144.5	7.2	1616	4	O15054

17	142	7.0	722	13	Q8UWU2	Q8UWU2 xenopus lae
18	138.5	6.9	981	15	Q92809	Q92809 abelson mur
19	135	6.7	559	11	Q91252	Q91252 mus musculus
20	134.5	6.7	662	5	Q9N3S5	Q9N3S5 caenorhabdit
21	132	6.5	112	11	P70305	P70305 mus musculus
22	132	6.5	764	5	Q24708	Q24708 drosophila
23	131.5	6.5	471	4	Q9UCX5	Q9UCX5 homo sapien
24	131.5	6.5	485	3	Q9UPP6	Q9UPP6 schizosacch
25	130.5	6.5	548	5	Q9YPU1	Q9YPU1 drosophila
26	129.5	6.4	594	4	Q92529	Q92529 homo sapien
27	129.5	6.4	594	4	Q8RAP2	Q8RAP2 homo sapien
28	128	6.3	17352	5	Q95YM2	Q95YM2 procambarus
29	126.5	6.3	728	4	Q9UPH9	Q9UPH9 homo sapien
30	126	6.2	377	11	Q55032	Q55032 mus musculus
31	126	6.2	377	13	P79956	P79956 xenopus lae
32	126	6.2	461	11	Q63789	Q63789 ratius norv
33	125.5	6.2	533	11	Q8VEN5	Q8VEN5 mus musculus
34	125.5	6.2	1258	11	Q8R457	Q8R457 mus musculus
35	125	6.2	269	11	Q9CT18	Q9CT18 mus musculus
36	125	6.2	377	11	Q92279	Q92279 mus musculus
37	125	6.2	377	11	Q99M51	Q99M51 mus musculus
38	125	6.2	460	4	Q96HD7	Q96HD7 homo sapien
39	125	6.2	595	13	Q92124	Q92124 xenopus lae
40	125	6.2	597	11	Q64509	Q64509 mus musculus
41	125	6.2	823	6	Q9TTY2	Q9TTY2 canis famli
42	124.5	6.2	217	11	Q9CX99	Q9CX99 mus musculus
43	124	6.1	410	5	Q24218	Q24218 drosophila
44	123.5	6.1	863	4	Q9UPV4	Q9UPV4 homo sapien
45	123	6.1	927	4	Q9UDE4	Q9UDE4 homo sapien

ALIGNMENTS

RESULT 1

Q9P2U9 PRELIMINARY; PRT; 376 AA.
ID Q9P2U9
AC Q9P2U9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MIST (Fragment).
GN MIST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20208989; PubMed=10744659;
RA Gotsuka R., Kanazashi H., Sasamura H., Fujimura Y., Hidaka Y.,
Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clnk involved in Ige
receptor-mediated mast cell degranulation.";
RL Int. Immunol. 12:573-580(2000).
DR EMBL: AB032369; BAA96241.1; --
DR HSSP: P23727; ZPNE.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
FT NON_TER 1
FT NON_TER 376
FT NON_TER 376
SQ SEQUENCE 376 AA; 43542 MW; 389421B629B028B2 CRC64;

Query Match 100.0%; Score 2020; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.6e-158;
Matches 376; Conservative 0; Mismatches 0; Gaps 0;
1 FQNFSLPKRNSWPRINSATGQYGRNKKPLDIERNFVAALDGAKSHSDDDYDPELRMBE 60
|||||

```

Db      1  FQNFSLPKNRSPRINSATGOYQRMNKPILDMERNFAVLGAKGSHDDYDDPELMBE 60
Qy      61  TWOSIKILPAPRIKESAYADTHFKVAMDTPLDTRTISIGOPTWNTQTLERVDKPI 120
Db      61  TWOSIKILPAPRIKESAYADTHFKVAMDTPLDTRTISIGOPTWNTQTLERVDKPI 120
Qy      121  SRDVRSONIKGASVAKNNIPLPPRPPLTLTKKQPLPEPESSRPPLSQHRTPEYOG 180
Db      121  SRDVRSONIKGASVAKNNIPLPPRPPLTLTKKQPLPEPESSRPPLSQHRTPEYOG 180
Qy      181  MFSQISLRLDSEVLEAEKVPNHQRPPESTHLENNQTOEIPLAISSSFTHSNHVSQVRD 240
Db      181  MFSQISLRLDSEVLEAEKVPNHQRPPESTHLENNQTOEIPLAISSSFTHSNHVSQVRD 240
Qy      241  HGGGMOPCSPQRCOPPCASCPHENILPYKYTSWRRPFRKDRKDVQHNEWYIGEYSRQA 300
Db      241  HGGGMOPCSPQRCOPPCASCPHENILPYKYTSWRRPFRKDRKDVQHNEWYIGEYSRQA 300
Qy      301  VEEAFMKENKDGSLVRDCSTKSKPEPVLAIFYENKYNKIRPLENOQFALGTGLRG 360
Db      301  VEEAFMKENKDGSLVRDCSTKSKPEPVLAIFYENKYNKIRPLENOQFALGTGLRG 360
Qy      361  DEKFDVSDIIEHYKN 376
Db      361  DEKFDVSDIIEHYKN 376

```

RESULT 2

```

09JMJ3  PRELIMINARY; PRT; 435 AA.
AC 09JMJ3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MIST.
GN CLNK OR MIST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20208989; PubMed=10744659;
RA Goltzuka R., Kanazashi H., Sasayama H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clnk involved in IgE
RT receptor-mediated mast cell degranulation.";
RL Int. Immunol. 12:573-580(2000).
DR EMBL: AB021220; BAA96240.1; -
DR HSSP: P23727; 2PNB.
DR MGD: MGI:1351468; Clnk.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD00093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 435 AA; 49513 MW; 4D1B03E2F0C61ED6 CRC64;

```

Query Match 56.0%; Score 1132; DB 11; Length 435;
Best Local Similarity 60.9%; Pred. No. 3.2e-85;

Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

```

Qy      1  FQNFSLPKNRSPRINSATGOYQRMNKPILDMERNFAVLGAKGSHDDYDDPELMBE 60
Db      19  FQNFSLPKNRSPRINSATGOYQRMNKPILDMERNFAVLGAKGSHDDYDDPELMBE 60
Qy      61  TWOSIKILPAPRIKESAYADTHFKVAMDTPLDTRTISIGOPTWNTQ-TLERYVDKPI 119
Db      79  AWPMSKILPAPRIKESAYADTHFKVAMDTPLDTRTISIGOPTWNTQ-TLERYVDKPI 119
Qy      120  ISRDVRSONIKGASVAKNNIPLPPRPPLTLTKKQPLPEPESSRPPLSQHRTPEY 178
Db      120  ISRDVRSONIKGASVAKNNIPLPPRPPLTLTKKQPLPEPESSRPPLSQHRTPEY 178

```

```

Db      139  TFKDVRSONIKGASVAKNNIPLPPRPPLTLTKKQPLPEPESSRPPLSQHRTPEY 198
Qy      179  QGMPQSILRLDSEVLEAEKVPNHQRPPESTHLENNQTOEIPLAISSSFTHSNHVSQVRD 238
Db      199  QGMPQSILRLDSEVLEAEKVPNHQRPPESTHLENNQTOEIPLAISSSFTHSNHVSQVRD 238
Qy      239  RDHGGMOPCSPQRCOPPCASCPHENILPYKYTSWRRPFRKDRKDVQHNEWYIGEYSR 298
Db      239  RDHGGMOPCSPQRCOPPCASCPHENILPYKYTSWRRPFRKDRKDVQHNEWYIGEYSR 298
Qy      299  QAVEAFMKENKDGSLVRDCSTKSKPEPVLAIFYENKYNKIRPLENOQFALGTGL 358
Db      317  QAVEAFMKENKDGSLVRDCSTKSKPEPVLAIFYENKYNKIRPLENOQFALGTGL 358
Qy      359  RGDEKFDVSDIIEHY 374
Db      377  RGDEKFDVSDIIEHY 392

```

RESULT 3

```

09QZE2  PRELIMINARY; PRT; 435 AA.
AC 09QZE2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CLNK.
GN CLNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=20029811; PubMed=10562326;
RA Cao M.Y., Davidson D., Yu J., Latour S., Veillette A.;
RT "Clnk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
RT stimulated Hemopoietic Cells ";
RL J. Exp. Med. 190:1527-1534(1999).
DR EMBL: AF187819; AAF14299.1; -
DR HSSP: P23727; 2PNB.
DR MGD: MGI:1351468; Clnk.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD00093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 435 AA; 49492 MW; 5CD27EC971FC0EAS CRC64;

```

Query Match 56.0%; Score 1132; DB 11; Length 435;
Best Local Similarity 60.9%; Pred. No. 3.2e-85;

Matches 229; Conservative 41; Mismatches 102; Indels 4; Gaps 3;

```

Qy      1  FQNFSLPKNRSPRINSATGOYQRMNKPILDMERNFAVLGAKGSHDDYDDPELMBE 60
Db      19  FQNFSLPKNRSPRINSATGOYQRMNKPILDMERNFAVLGAKGSHDDYDDPELMBE 60
Qy      61  TWOSIKILPAPRIKESAYADTHFKVAMDTPLDTRTISIGOPTWNTQ-TLERYVDKPI 119
Db      79  AWPMSKILPAPRIKESAYADTHFKVAMDTPLDTRTISIGOPTWNTQ-TLERYVDKPI 119
Qy      120  ISRDVRSONIKGASVAKNNIPLPPRPPLTLTKKQPLPEPESSRPPLSQHRTPEY 178
Db      139  TFKDVRSONIKGASVAKNNIPLPPRPPLTLTKKQPLPEPESSRPPLSQHRTPEY 198
Qy      179  QGMPQSILRLDSEVLEAEKVPNHQRPPESTHLENNQTOEIPLAISSSFTHSNHVSQVRD 238
Db      199  QGMPQSILRLDSEVLEAEKVPNHQRPPESTHLENNQTOEIPLAISSSFTHSNHVSQVRD 238
Qy      239  RDHGGMOPCSPQRCOPPCASCPHENILPYKYTSWRRPFRKDRKDVQHNEWYIGEYSR 298
Db      239  RDHGGMOPCSPQRCOPPCASCPHENILPYKYTSWRRPFRKDRKDVQHNEWYIGEYSR 298

```

Db 259 RDHTGSMOHCBAORCAASHSP--RMLPYENTNSEKPDPTKPKDEKDYQWQNEWYIGEXSR 316
 Qy 299 QAVEBAFMKKNKSGFLVRCDSYTSKEEPPYLAIFYENKYNVKTRELERNOQFALGTGL 358
 Db 317 QAVEBVLKKNKKGFTLVRCDSYTSKAEPPYLVYFNGKYNVKTRELERNOQFALGTGL 376
 Qy 359 RGDEKDSVEDIIEHY 374
 Db 377 RGNEMFDSVEDIIEHY 392

RESULT 4

09YGC1 PRELIMINARY: PRT: 552 AA.
 ID 09YGC1
 AC 09YGC1
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE B cell linker protein BLNK.
 GN BLNK OR BASH.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99146381; PubMed=10023776;
 RA Ishai M., Kurosaki M., Pappu R., Okawa K., Ronko I., Fu C.,
 RA Shihata M., Iwamatsu A., Chan A.C., Kurosaki T.;
 RT "BLNK required for coupling Syk to PLC gamma 2 and Rac1-JNK in B
 cells";
 RL Immunity 10:117-125(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99049801; PubMed=9834055;
 RA Gotsuka R., Fujimura Y., Mamada H., Umeda A., Morimura T.,
 RA Uetsuka K., Doi K., Tsuji S., Kitamura D.;
 RT "BASH, a novel signaling molecule preferentially expressed in B cells
 of the bursa of Fabricius";
 RL J. Immunol. 161:5804-5808(1998).
 RL EMBL; AF089727; AAD12783.1; -;
 DR HSSP; P08487; 2PID.
 DR HSSP; P08487; BAA36275.1; -;
 DR InterPro; IPR002965; P_Rich_extensn.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PR01217; PRICEXTENSN.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS0001; SH2; 1.
 DR PROSITE; PS5001; SH2; 1.
 SQ SEQUENCE 552 AA; 61822 MW; FB232179BE38D072 CRC64;

Query Match 16.5%; Score 334; DB 13; Length 552;
 Best Local Similarity 24.3%; Pred. No. 2,9e-19;
 Matches 112; Conservative 60; Mismatches 157; Indels 132; Gaps 15;

Qy 32 WEERNAVLIDGAKGHS-----DDYDDELMEETWOSIKILPARPIRESY 78
 Db 83 WSDDFSDENPDGHSDSEMYVSEENPDSDYEPPESEQEK-----KIPSPISRGEX 138
 Qy 79 AD---HYKRYVMDPLPLDTRSI-----SIGOPTWMTQFTL-----ERDKPI 120
 Db 139 ADKRTSHHQLPTNKPPLSPSSALPRPKPSPSPAKPKPLAKRECSDEEDNTIVPY 198
 Qy 121 SRDVSQNIKGASV-----RKNKIPLPRLPILPKKYQLPPEPSS 165
 Db 199 DNDNDVYIPTESSTPPAKPVPVNRMKPPAKGALPTPPKPLASDMQEVYEPPEEEL 258
 Qy 166 RPEPLSQHTEPEYQMP-----QISLDLSEVLEAEK-----VPHNQKPESTHL 211
 Db 259 SPPPVTRFTKP---LPATRAQNAEHSNHSMTRESPKLDASRNILPLPNRIAPKTDHE 314

Qy 212 LEN-----QNTQE-----IPLATSSSS-----FTSNHVSQN 238
 Db 315 ANNDENHSFNTQESKPPGCAADSPDLRALKTSNAVPAKPCPLPSMDTFTVEDKETA 374
 Qy 239 RDHRGGMQ---PCSPORPCAPPAS-----GSPHENILPYKYSWRP 275
 Db 375 ADNRGSSHEFFPLPLPGTPEKSSLOKPLVLPKVEAPSRALGTPSPHSSISSSTA--- 431
 Qy 276 PPKRSRDKVOHNEWYIGEYSROAVEBAFMKKNKSGFLVRCDSYTSKEEPPYLAIFYE 335
 Db 432 ----DQAGYASKAWYATCDKRYAEDALYRSNKGSGFLIRKSSGQDSROPYTLVYFN 486
 Qy 336 NKVYVVKIRPLERNOQFALGTGLRGDEKDSVEDIIEHYKN 376
 Db 487 RKNVNIPIRFESTRQYALGRKCGEENRDSVAELVENHQH 527

RESULT 5

09DG07 PRELIMINARY: PRT: 530 AA.
 ID 09DG07
 AC 09DG07
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE SLP-76 adaptor protein.
 GN SLP76.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20451093; PubMed=1093915;
 RA Ishai M., Kurosaki M., Inabe K., Chan A.C., Suganuma K., Kurosaki T.;
 RT "Involvement of LAT, Gads, and Grb2 in Complementatation of SLP-76 to
 the Plasma Membrane";
 RL J. Exp. Med. 192:847-856(2000).
 RL EMBL; AF226986; AAG18493.1; -;
 DR HSSP; P12931; 1SHD.
 DR HSSP; P12931; SAM.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS5001; SH2; 1.
 SQ SEQUENCE 530 AA; 59580 MW; CF6CBCEBD79909A CRC64;

Query Match 15.7%; Score 317.5; DB 13; Length 530;
 Best Local Similarity 26.9%; Pred. No. 6,2e-18;
 Matches 105; Conservative 53; Mismatches 145; Indels 87; Gaps 14;

Qy 46 HSDDDYDDELMEETWOSIKILPARPI-KESFYAD-----THYFKYAMD----- 89
 Db 140 HSDGYEPPPSNNDENAHNV-IFPAKSLANNNDYIDRPTSSSHQPPVPPQRPGRSPAP 198
 Qy 90 -----TPLPLDTRTSISIGP-----TWNTQRLERVDKPISRDVSQNIKG 131
 Db 199 ASFGCGASLPAPFPLPGNNDRVKSKPPAPSIDRSTKPLDRLGPPERE---NPVPG 255
 Qy 132 DASVRKNKILPLPPR---PLTLPKYQLPPEPSSR---PLLSQKHTPEYQMGMS 183
 Db 256 RKPGEPEKLTPLQALGSQLAMMP-----PVPPSDRYERGNPPLKQILPVKQGMQ 310
 Qy 184 QISLRDSEVLEAEKYPHNOQKPESTHLENON-----QIEIPLATSSSFTSNHS 235
 Db 311 QKRPDEEDHIIPORAVPQISLPYSSNTTPSKIKAPKPGSGSIFGASARLSMS----- 367
 Qy 236 VONRDRGGMQPCSPORCOPPASCSPHENILPYKYTSWRPPFP-----KSRDKRD 285

```

Db 368 -----SLRPRLPGLNN--SRSPSRGT-----ADLRPLPISROTAHOTNTEDEED 412
OY 286 VOHNEMYIGESYRQAVVEAFMKENKDGSLVLDCKSTKKEEYVLAAYENKVVVVKIRF 345
DB 413 SLNDEMVAAYISRPAAEAALRKINKDGTFLVDSRSKRTTHRYLMLVLYKDKVYNIQIRY 472
OY 346 LERNOQFALGTGLRGDKDFDSVEDIEHYK 375
DB 473 QEODOTYLLGTGLKGEDEFFSVAHIIIDYEQ 502

```

RESULT 6

```

O920L0 PRELIMINARY: PRT: 525 AA.
AC O920L0:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE SLP-76 adaptor protein.
GN SLP76.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fujii Y., Golthuka G., Wakahara S., Hara T., Nakao T.;
RT "Molecular cloning of SLP-76, a 76-kDa adaptor protein in RBL-2H3.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB072980; BAB71779.1;
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRODOM: PD000093; SH2; 1.
DR PROSITE: PSS0001; SH2; 1.
SQ SEQUENCE 525 AA; 59521 MW; 0BICDD40DABE7372 CRC64;

```

```

Query Match 15.78; Score 316.5; DB 11; Length 525;
Best Local Similarity 25.98; Pred. No. 7,4e-18;
Matches 11; Conservative 49; Mismatches 112; Indels 173; Gaps 19;

```

```

OY 20 GOYORMNKPLDWMERNFAAVIDGAKGSHDDYDDELMEETWQSIILPARPIKESEYA 79
DB 126 GDYSPN-----EEELVDDAA-----DYEPSPNDEALON-SILPRAKRP----- 166
OY 80 DTHFKVAMDTPLDLTRTISISIGOPTWNTQTRLEVDKPRISRDVSQNTKGDASVKNK 139
DB 167 -----PMTNSK-----YIDRPTGKVSQP----- 186
OY 140 IPLRPRLITLPKKYQRLP-----PEP-----ESSRPLSQ-- 171
DB 187 -PVRPQRPMAALP---PLPTGRNHSNHEPSPNNRNHKTAKRAPSIDRSTKRPRLDSL 241
OY 172 -----NHTPEYQ--GMPQSILRLDSE---VLEAK--VPH 201
DB 242 APRLDREPLIGKPKSDKSPAPLGRHRLPKIQKRLPRAMORHENEIGLTLAKRPVPR 301
OY 202 NQKRP-----STHLENOQTQ-----EPLAISSSSF 229
DB 302 HGRPRDRONDEDVHQRLPHPSLPSMSNTFSPRSKTPKNTFPLPHMGAISESNI 361
OY 230 TTSNHSVQNDHNGMOPCS--PORCOPRASCSPHEN---ILPYKYTSMRPPPKRSDK 284
DB 362 -----GFOQASALPSYFSGSPSSRPVRNREVNANLPLVPN-RPQPSGEE 407
OY 285 DVQHNEMYIGESYRQAVVEAFMKENKDGSLVLDCKSTKKEEYVLAAYENKVVVVKIR 344
DB 408 SPLDEEYVYSITRPEAEALRKINQDGTFLVDSRSKRTYNNPYLVLYLTKDKVYNIQIR 467
OY 345 FLENOQFALGTGLRGDKDFDSVEDIEHYK 375
DB 468 YQESQVYLLGTGLKGEDEFLVSIIIDYFR 498

```

RESULT 7

```

O922M0 PRELIMINARY: PRT: 533 AA.
AC O922M0:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lymphocyte cytosolic protein 2.
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC006948; AAH06948.1;
DR MGD: MGI:1321402; Lcp2.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRODOM: PD000093; SH2; 1.
DR PROSITE: PSS0001; SH2; 1.
SQ SEQUENCE 533 AA; 60238 MW; 50AEA025EF0DAD01 CRC64;

```

```

Query Match 15.68; Score 314.5; DB 11; Length 533;
Best Local Similarity 30.08; Pred. No. 1.1e-17;
Matches 121; Conservative 44; Mismatches 169; Indels 69; Gaps 18;

```

```

OY 20 GOYORMNKPLDWMERNFAAVIDGAKGSHDDYDDELMEETWQSIILPARPI--KESE 77
DB 126 GDYSPN-----EEELVDDAA-----DYEPSPNDEALON-SILPNSFHTNSM 172
OY 78 YAD-TTHFKVAMDTPL-DLTRTISISIGOPTWNTQTRLEVDKPRISRDVSQNTKGDASY 135
DB 173 YIDRPTGKVSQDPVPRPLPRAALP-PLPTGRNHSPLSPHNPHEPSSGNNK----- 226
OY 136 RKNKIRP-----PRPLITLPKKYQRLP-----PEPSSRPLSQRTPEYQ--GM 181
DB 227 -TAKLPAPSIDRSTKRPPLDSLALDLREPTLIGKPKSPKAPLGRH-LRKIQKPL 284
OY 182 PSQISLRLDSEVL-----EAEKYPHNRKPESTHLE--NQNTQEIPLAISSSSF 233
DB 285 PRAMDREHNERLGRPVTRKRPVPRHGRGPRDREDDVHQRLPQSLPSMSNTFPS 344
OY 234 HSYO-----NRDHRGMOPCS--PORCOPRASCSP---HNTILPYKYTS 272
DB 345 RSYQPSKSKNTPLAHMGAISESNTGFOQASALPSYFSQPGNRPLRSEGRNLPLVPN 404
OY 273 WRPFPRKSDKDVQHNEMYIGESYRQAVVEAFMKENKDGSLVLDCKSTKKEEYVLAAY 332
DB 405 -RPQPSGEEETRLDEEWVYSITRPEAEALRKINQDGTFLVDSRSKRTYNNPYLVLY 463
OY 333 FYENKYYNKKIRFLERNOQFALGTGLRGDKDFDSVEDIEHYK 375
DB 464 LYKDKVYNIQIRYQESQVYLLGTGLKGEDEFLVSIIIDYFR 506

```

RESULT 8

```

O75499 PRELIMINARY: PRT: 433 AA.
AC O75499:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE B cell linker protein BLNK-S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

RX MEDLINE=98001722; PubMed=9341187;
 RA Fu C., Chan A.C.;
 RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
 RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
 RT receptor activation.";
 RL J. Biol. Chem. 272:27362-27368(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98361304; PubMed=9697839;
 RA Fu C., Turk C.W., Kurossaki T., Chan A.C.;
 RT "BINK: a central linker protein in B cell activation.";
 RL Immunity 9:93-103(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20050956; PubMed=10583958;
 RA Mingeshtl Y., Kohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
 RA Campana D., Chan A.C., Conley M.E.;
 RT "An essential role for BINK in human B cell development.";
 RL Science 286:1954-1957(1999).
 DR EMBL; AF068181; AAC39937.1;
 DR EMBL; AF180756; AAF20383.1;
 DR EMBL; AF180740; AAF20383.1; JOINED.
 DR EMBL; AF180741; AAF20383.1; JOINED.
 DR EMBL; AF180742; AAF20383.1; JOINED.
 DR EMBL; AF180743; AAF20383.1; JOINED.
 DR EMBL; AF180744; AAF20383.1; JOINED.
 DR EMBL; AF180745; AAF20383.1; JOINED.
 DR EMBL; AF180746; AAF20383.1; JOINED.
 DR EMBL; AF180748; AAF20383.1; JOINED.
 DR EMBL; AF180749; AAF20383.1; JOINED.
 DR EMBL; AF180750; AAF20383.1; JOINED.
 DR EMBL; AF180751; AAF20383.1; JOINED.
 DR EMBL; AF180752; AAF20383.1; JOINED.
 DR EMBL; AF180753; AAF20383.1; JOINED.
 DR EMBL; AF180754; AAF20383.1; JOINED.
 DR EMBL; AF180755; AAF20383.1; JOINED.
 DR HSSP; P29354; 1BMB.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS50001; SH2; 1.
 SQ SEQUENCE 433 AA; 48229 MW; 0B36FE9FCF5DC7DC CRC64;
 Query Match 14.3%; Score 289.5; DB 4; Length 433;
 Best Local Similarity 24.6%; Pred. No. 9,7e-16;
 Matches 93; Conservative 57; Mismatches 127; Indels 101; Gaps 15;
 QY 43 AKGSHDDYDDDELMEETWOSIKILPAPRIKESYADTHYKAVMDTLPIDTSTISI 102
 DB 88 AENADDSYEPPVEOE-----TRPVHLPALFARGEYIDNR-----SSQRHSPFSKTLPS- 138
 QY 103 GQPTWNTQRLERVDKPISRDVRSNLIKGDASVRKKKIPLPERRPLI----- 149
 DB 139 -RPSW-----PSEKARLTSLPALTLAQKQPV-PRKKGLEDEADYVVEVDN 185
 QY 150 ----TLPKYQPLPPEP-----ESSRPLSORHTFPPEYQMPQ-----ISLRDIS 191
 DB 186 DENYIHPTESSPPPEKGRNSGAMETKSPPPAPSPPLPRAGKKPTPLKTPVVASQQMS 245
 QY 192 EYLEAEKVHNOCKPESTHLENOQTQELPLATSSSFTTSNHSYVNRDHRGMOQCSQP 251
 DB 246 SVCEERPIP-AERHGRSSHROE-----AVQSPVF----- 273
 QY 252 RCPAPASCSPHENILPY-KYT-SWRP-----PPKRSRDKDVQHN-----WYIGESR 298
 DB 274 ---PPAQKHQKPIPLPRTGEGNPTVGGPLPSFSSNSTISQEGVLCKPMYAGACDR 330
 QY 299 QAVEEAFMKNKDGSLVRDCSTKSKKEEPVLAVFYENKYNKIFLERNOQFALGTGL 358
 DB 331 KSAEELHRSNKDGSFLIRKSSGSHDSKQPTLVVFFNKRKYNIPIVFIEATKOYALGRKK 390
 QY 359 RDEKFDSDVEDLIEHYKN 376

DB 391 NGEYFGSAELIRNHQH 408
 RESULT 9
 Q8WV28
 ID 08WV28 PRELIMINARY; PRT; 456 AA.
 AC 08WV28;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE B-cell linker.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TONSIL;
 RA Strausberg R.;
 RL Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC018906; AAH18906.1;
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS50001; SH2; 1.
 SQ SEQUENCE 456 AA; 50465 MW; ED6D424A035D1792 CRC64;
 Query Match 14.3%; Score 289; DB 4; Length 456;
 Best Local Similarity 25.6%; Pred. No. 1.1e-15;
 Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;
 QY 43 AKGSHDDYDDDELMEETWOSIKILPAPRIKESYADTHYKAVMDTLPIDTSTISI 102
 DB 88 AENADDSYEPPVEOE-----TRPVHLPALFARGEY-----IDNRSS- 125
 QY 103 GQPTWNTQRLERVDKPISRDVRSNLIKGDASVRKKKIPLPERRPLIPLPKYQPLPPEP 162
 DB 126 -----ORHSPFSKTLPSK-----PSWPSKALITSLPALTLAQKQV-PPRP 168
 QY 163 -----ESSRP-----PLSORHTFPPEYQ-----GMPQSISLSD 189
 DB 169 KGLEDEADYVVEVDNDENYIHPTESSPPPEKAPMVRSTKPNSTPASPGRASGN 228
 QY 190 LSEYLEAEK-----VPHNQRK-----ESTHLENOQTQELPLATSSSFTTSNHSVQ 237
 DB 229 -SGAMETKSPPPAPSPPLPRAGKKPTPLKTPVVASQN-----ASSVCEERPIP 277
 QY 238 NRDRHG--MOPCSPQRCQRPASCSPHENILPY-KYT-SWRP-----PPKRSRDKDVQHN 289
 DB 278 AERHGRSSHROEAVOSPVPPAQKHQKPIPLPRTGEGNPTVGGPLPSFSSNSTISQ 337
 QY 290 E-----WYIGESRQAVEEAFMKNKDGSLVRDCSTKSKKEEPVLAVFYENKYNK 342
 DB 338 EAGVLCKPMYAGACDRKSAEELHRSNKDGSFLIRKSSGSHDSKQPTLVVFFNKRKYNIPI 397
 QY 343 IRLERNOQFALGTGLGDEKFDSDVEDLIEHYKN 376
 DB 398 VRIEATKOYALGRKKNGEYFGSAELIRNHQH 431
 RESULT 10
 ID 075498 PRELIMINARY; PRT; 456 AA.
 AC 075498;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE B cell linker protein BINK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98001722; PubMed=9341187;
 RA Fu C., Chan A.C.;
 RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
 RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
 RT receptor activation."
 RL J. Biol. Chem. 272:27362-27368(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98361304; PubMed=9697839;
 RA Fu C., Turk C.W., Kurosaki T., Chan A.C.;
 RT "BLNK: a central linker protein in B cell activation."
 RT Immunity 9:93-103(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20050956; PubMed=10583958;
 RA Mueglisthal Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
 RA Campana D., Chan A.C., Conley M.E.;
 RT "An essential role for BLNK in human B cell development."
 RT Science 286:1954-1957(1999).
 RL [1]
 DR EMBL: AF068180; AAC3936.1;
 DR EMBL: AF180756; AAF20382.1;
 DR EMBL: AF180740; AAF20382.1; JOINED.
 DR EMBL: AF180741; AAF20382.1; JOINED.
 DR EMBL: AF180742; AAF20382.1; JOINED.
 DR EMBL: AF180743; AAF20382.1; JOINED.
 DR EMBL: AF180744; AAF20382.1; JOINED.
 DR EMBL: AF180745; AAF20382.1; JOINED.
 DR EMBL: AF180746; AAF20382.1; JOINED.
 DR EMBL: AF180747; AAF20382.1; JOINED.
 DR EMBL: AF180748; AAF20382.1; JOINED.
 DR EMBL: AF180749; AAF20382.1; JOINED.
 DR EMBL: AF180750; AAF20382.1; JOINED.
 DR EMBL: AF180751; AAF20382.1; JOINED.
 DR EMBL: AF180752; AAF20382.1; JOINED.
 DR EMBL: AF180753; AAF20382.1; JOINED.
 DR EMBL: AF180754; AAF20382.1; JOINED.
 DR EMBL: AF180755; AAF20382.1; JOINED.
 DR HSSP: P29354; 1BM8.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 SQ SEQUENCE 456 AA; 50466 MW; 95F1D5485D0D397 CRC64;
 Query Match 14.3%; Score 289; DB 4; Length 456;
 Best Local Similarity 25.6%; Pred. No. 1,1e-15;
 Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;
 OY 43 AKGSHDDYDPELMEETWOSIKILPARPIKESYADTHYKAVAMDPLDRTSISI 102
 DB 88 AENADDSYEPPEVE-----TRPVHPLPFAFGY-----IDNRSS--- 135
 OY 103 GQPMWNTQTLERVDKRTSDVNSQNTKGASVRKKNIPPLPPPLTLTKKKYQPLPEP 162
 DB 126 -----QRHSPFSKTLPSK-----PSWPESEKARLTSLPALALQKQ-VPPKP 168
 OY 163 -----ESSRP-----PLSQHRTPEVQ-----GMSQSLSLRD 189
 DB 169 KGLLEDEADYVVEEDNDENKIHPTSSSPPEKAPVNRSTPNSTPASPCTAGSRN 228
 OY 190 LSEVLEAEK-----VPHNQRK-----ESTHLEENQNTQETPLAISSSEFTSNHSVQ 237
 DB 229 -SAGMETKSPPPAPSPPLPRAGKKPTPLKTTVAASQ-----ASSVCEKRP 277
 OY 238 NRDRHG--WQPSQRCQPCPASCSPHENLPI-KYT-SMRP-----PPKRSRDKVQNH 289
 DB 278 AERHRRSSHRQEAQVQSPVPPAPKQIHKRPIPLRFTGEGNPVDDGLPFSFSSNSTISQ 337
 OY 290 E-----WYIGETSRQAVAEAFMKENKDGSLVYRDCSTKSKPEPYLVAVFYENKYYNKK 342

DB 338 EAGVLCRWYAGACDRSAEALHRSNKDGSEFLIRKSGHDSKQPYTLVFFNKKRYNIP 397
 OY 343 IRFLERNOEAFALGTGLNGDEKEDSVEDIIEHYKN 376
 DB 398 VRFTEATKQVALGKRKKNGEEFSGVAELINHHQ 431
 RESULT 11
 ID O9QUN3 PRELIMINARY; PRT; 457 AA.
 AC O9QUN3;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE SH2-containing leukocyte protein 65 (lymphocyte antigen 57).
 GN SLP-65 OR BASH OR LY57.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-LYMPHOID;
 RX MEDLINE=98372771; PubMed=9705962;
 RA Wienands J., Schweikert J., Wollischled B., Jumaa H., Nielsen P.J.,
 RA Reith M.;
 RT "SLP-65: A new signalling component in B lymphocytes which requires
 RT expression of the antigen receptor for phosphorylation."
 RL J. Exp. Med. 188:791-795(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-LYMPHOID;
 RA Wienands J., Larbolette O., Reith M.;
 RT "Evidence for a preformed transducer complex organized by the B cell
 RT antigen receptor."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:7865-7870(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Okamoto N., Hayashi K., Tsuji S., Gotsuka R., Kitamura D.;
 RT "BASH: B lymphocyte adaptor protein containing SH2 domain."
 RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON-ETN;
 RA Nielsen P.J., Guenet J.L.;
 RT "The murine SLP-65 gene."
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: Y17159; CAA7666.1;
 DR EMBL: AB015290; BAA34944.1;
 DR EMBL: AJ298054; CAC18565.1;
 DR HSSP: P23727; 1BFT.
 DR MGD: MGT:96876; LY57.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 KW B-cell.
 FT CONFLICT 133 133 S -> N (IN REF. 2).
 FT CONFLICT 147 147 A -> T (IN REF. 2).
 FT CONFLICT 148 150 RLA -> GIG (IN REF. 2).
 FT CONFLICT 197 198 PP -> AT (IN REF. 2).
 FT CONFLICT 391 391 N -> K (IN REF. 2).
 FT CONFLICT 444 446 TRD -> SKH (IN REF. 2).
 SQ SEQUENCE 457 AA; 50670 MW; 66C93D4FDDF9D260 CRC64;
 Query Match 14.0%; Score 282; DB 11; Length 457;
 Best Local Similarity 23.8%; Pred. No. 4.3e-15;
 Matches 96; Conservative 58; Mismatches 157; Indels 92; Gaps 15;
 OY 32 WERNFAAVLDGAKGHS-----DDDYDPELMEETWOSIKILPARPIKESY 78

```

DB 64 MSDDSDSYENDEHSEDEMYVMPAEETGDDSEYEPPEAQ-----QTRVHPALPPTRGEX 119
OY 79 ADHYFKYAMDPTPLDTRTSTISIGOPTWNTOTRLERVDKPLSRDVRSONIKGDA5VRKN 138
DB 120 VDNR-----SSQHRSPSPFKTLPSS--KPSW-----PSAKARLASTLPAPNSLOKP 162
OY 139 KIPPLPPPLI-----TLPKYOPLPPEP-----ESSRPLSLORHPP 176
DB 163 QVP-FKPKDLEDEADYVPEVDNDENYIHPRESSPPPAEKAPMVNRSTKPPSSSKHNSP 221
OY 177 --EVOGMSQISLRDLSEVLEAEKVPHNQKPESTHLENQTOELPLAISSSFTTNSH 234
DB 222 PCTVAGRNSGVWDSKSSLPAPSPPLPRAGKP-ATPL-----KTTVP-PLPANSVCEK 275
OY 235 SVONDRHRCG--MOPCSPORCOPPASCSPHENILPKYKTSWMPPEPK-----279
DB 276 PVPARHRSSSHRODTVOSPVPPTQKPVHQKVPVL-----PRPEAGSPADGPFHSF 329
OY 280 -----RSRKRQVQHNEMWIGEYSRQAVEAEAFMKENKDGSLVROCSKSEEPYLAVF 333
DB 330 PENSTFADQEAELGLGKPMWAGACDRKSAEALHRSNKDGSFLIRKSSFGHDSKQPYTLVAF 389
OY 334 YENKYVNVKIRFLERNQOFALGTGLRGDEKFDSDVEDIEHYKN 376
DB 390 FKKRYVNIPIRFEATKQYALGKKNGEYFGSVVEIVNSHOH 432

RESULT 12
O88504 PRELIMINARY: PRT: 457 AA.
ID 088504:
AC 088504:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE B cell linker protein BLNK.
GN LY57.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Gbd2, and Yav after B cell antigen
RT receptor activation.*
RL J. Biol. Chem. 272:27362-27368(1997).
RN 12]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turk C.W., Kurosaki T., Chan A.C.;
RT "BLNK: A Central Linker Protein in B Cell Activation."
RL Immunity 9:93-103(1998).
DR EMBL: AF068182; AAC40206.1;
DR HSSP: P23727; 1BFT.
DR MGD: MGI:96878; LY57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 457 AA; 50803 MW; 66D235796A6C45F0 CRC64;

Query Match 13.5%; Score 273; DB 11; Length 457;
Best local similarity 23.8%; Pred. No. 2,4e-14;
Matches 96; Conservative 60; Mismatches 155; Indels 92; Gaps 16;

OY 32 WEKNFAVLDGAKGHS-----DDVDDELPMHEETWQSIKILPAPRKSEY 78
DB 64 WSDFDSDYENDEHSEDEMYVMPAEETGDDSEYEPPEAQ-----QTRVHPALPPTRGEX 119

```

```

OY 79 ADHYFKYAMDPTPLDTRTSTISIGOPTWNTOTRLERVDKPLSRDVRSONIKGDA5VRKN 138
DB 120 VDNR-----SSQHRSPSPFKTLPSS--KPSW-----PSAKARLASTLPAPNSLOKP 162
OY 139 KIPPLPPPLI-----TLPKYOPLPPEP-----ESSRPLSLORHPP 176
DB 163 QVP-FKPKDLEDEADYVPEVDNDENYIHPRESSPPPAEKAPMVNRSTKPPSSSKHNSP 221
OY 177 --EVOGMSQISLRDLSEVLEAEKVPHNQKPESTHLENQTOELPLAISSSFTTNSH 234
DB 222 PCTVAGRNSGVWDSKSSLPAPSPPLPRAGKP-ATPL-----KTTVP-PLPANSVCEK 275
OY 235 SVONDRHRCG--MOPCSPORCOPPASCSPHENILPKYKTSWMPPEPK-----280
DB 276 PVPARHRSSSHRODTVOSPVPPTQKPVHQKVPVL-----PRPEAGSPADGPFHSF 329
OY 281 -----SDRK-DVQHNEMWIGEYSRQAVEAEAFMKENKDGSLVROCSKSEEPYLAVF 333
DB 330 PENLTFADQEGELGLGKPMWAGACDRKFAEALHRSNKDGSFLIRKSSFGHDSKQPYTLVAF 389
OY 334 YENKYVNVKIRFLERNQOFALGTGLRGDEKFDSDVEDIEHYKN 376
DB 390 FKKRYVNIPIRFEATKQYALGKKNGEYFGSVVEIVNSHOH 432

RESULT 13
O9D413 PRELIMINARY: PRT: 297 AA.
ID O9D413:
AC O9D413:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE 4933424C13RIK.
GN 4933424C13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Resole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK016886; BAB30480.1;
DR HSSP: P16277; 1BTK.
DR MGD: MGI:1918380; 4933424C13RIK.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 297 AA; 32015 MW; BECE6B8ACF70685C CRC64;

```

```

Query Match      8.8%; Score 177; DB 11; Length 297;
Best Local Similarity 25.6%; Pred. No. 1,1e-06;
Matches 68; Conservative 32; Mismatches 106; Indels 60; Gaps 9;

QY 138 NIIPLPPPPPLTLTKKYPPLPEPPSSRPPLSORHTPEVOGMSQ-----184
DB 37 SKPPLEPPPOSTMA-----RGLPINSFPPIRPTSGYH-FPLKATAMPQPAKQGVFGROGR 90
QY 185 -ISLR-----DLSEVLEAEKVPHNOKRPESTHLENNQOEPLAISS-----SSF 229
DB 91 GTSARNVTKKPPEDITYLEEPPDPVPLTRLSLKALI-----PPVPLPTSGLPKSVAGY 145
QY 230 TTSNHSVONRDRGNGQPCSPQPCSPHSPHENILPKYTSWRRPPPKRSRDKRDVOHN 289
DB 146 QEARCAGMDGALKAGRR-LSASSIAPALSSVAEN-----GSLIGQ 185
QY 290 EWTIGESVQAVEEAFPMKENKDGSPFLVRCSPKSKKEEPPVLAFFENKYNKIFLERN 349
DB 166 PWSGNCNRQSVERRALLHFKQKDAIKVRLSSGPHSSQPTTLAVLLRGVFNIPRIQLOGG 245
QY 350 QQFALG-TGLRGDEKFDVSDIIEHY 374
DB 246 HHYALGREGRNHEELFSSVAAMVOHY 271

RESULT 14
ID 054737 PRELIMINARY; PRT; 516 AA.
AC 054737;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE BCA protein (fragment).
GN Lys7 OR BCA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/C;
RA MEDLINE=98346794; PubMed=9683264;
RA Gangl-Peterson L., Peterson S., Shapiro L., Golding A., Carlachio R.,
RA Cohen D.I., Margulies D.H., Cohen P.L.;
RA "Bca -- An Activation-related B-cell Gene.";
RL MOL. Immunol. 35:55-63(1998).
DR EMBL; AJ222814; CAAL1002.1; -.
DR MGD; MGI:96878; Lys7.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2. 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
FT NON_TER 1
SQ SEQUENCE 516 AA; 57822 MW; F31E65A7DC876FDD CRC64;

Query Match      7.6%; Score 153; DB 11; Length 516;
Best Local Similarity 20.3%; Pred. No. 0.00022;
Matches 77; Conservative 46; Mismatches 130; Indels 126; Gaps 15;

QY 32 WERNFAVLDCAKGS-----DDYDDELMEETQSKILPARIKSEY 78
DB 98 WSDFDSDYEPDHSSEKVMPEETGDSYEPPEAQ-----QTRVVPALPFTRGEX 153
QY 79 ADHYFVAMDTPLDTRTSISIGPTWMTQTRLEVRDPISR--DVRSQNIKGDASVR 136
DB 154 -----VNRSS-----QTLRPSAHFVSPAGLQKAKGLG 184
QY 137 KKKPLP-----PPRPLTLTKKYQPLP-----EPSSRPPLSQR-----172
DB 185 ASTLPANSLQKQVPPKPKPOLLEADYVVPVEDNDENYIHPRESSASNLRLRLPMSID 244

```

```

QY 173 ---HMFPEV-----QGMPSQISL-----RDISEVLEAEKVP-----HN 202
DB 245 QPSQTPQSTCARSVDAGQVGGSTPSHLCLPHNPHYGLGSLDHLRLPLPYRMHQ 304
QY 203 QRKPESTHLE-NQNTQEIPLAISSSFTTSHSVONRDRGNGQPCSPQRCQPPASCSP 261
DB 305 MEVKKSLPLLSANRRSSHRQDTPVQSPVPPQKPYHOK-----PVLPKAGSPAAQCP 357
QY 262 HENILPYTTSWRRPPPKRSRDKVO---HNWTIGESVQAVEEAFPMKENKDGSPFLVR 317
DB 358 FHS-----FPNSTFADQEAELGLKPMYVAVPNRKSAAEALHRSNKDGSPFLR 405
QY 318 DOSTKKEEPPVLAFFEN 336
DB 406 KSSGHSQPTTLVAFFKS 424

RESULT 15
ID P70304 PRELIMINARY; PRT; 724 AA.
AC P70304;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Phosphoinositide 3-kinase p85alpha.
GN PIK3R1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN(5); PART OF CLONE);
RX MEDLINE=97079666; PubMed=8921377;
RA Friedman D.A., Cantley L.C., Carpenter C.L.;
RT "Structural organization and alternative splicing of the murine
RL phosphoinositide 3-kinase p85 alpha gene.";
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; U50413; AAC52847.1; -.
DR HSSP; P23727; 1BFI.
DR MGD; MGI:97583; PIK3R1.
DR InterPro; IPR001720; PI3kinase_P85.
DR InterPro; IPR00198; RhogAP.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00620; RhogAP; 1.
DR Pfam; PF00017; SH2; 2.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00678; PI3KINASEP85.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 2.
DR SMART; SM00324; RhogAP; 1.
DR SMART; SM00252; SH2; 2.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PSS0001; SH2; 1.
DR PROSITE; PSS0002; SH3; 1.
KW kinase; SH3 domain.
SQ SEQUENCE 724 AA; 83501 MW; BCC21211DA6FIC17 CRC64;

Query Match      7.3%; Score 147; DB 11; Length 724;
Best Local Similarity 21.7%; Pred. No. 0.0011;
Matches 76; Conservative 38; Mismatches 97; Indels 140; Gaps 17;

QY 141 PLPPRPLTLTKKYQPLPEPPSSR-----PLSQRHHTPEVOGMSQISLR 188
DB 85 PTKPRP---PR---PLVAPGSSKTEADTEQALPLPLAQCFAPPDV-APPLIKILL 136
QY 189 DLSEV-LEAEKVPHNOKRPESTHLE-----NQNTQEIPLAISSSFTTSHSVONR 239
DB 137 EALEKGLSESTLYKTOSSNPALRLQLDCAVASVLEKIDVHVLADAKRYLADPN- 195
QY 240 DHRGNGQPCSP-----ORCOPPASC-----SPHENILPYKY-----270

```



```

Db 196 -----P|P|V|A|V|N|E|M|S|L|A|Q|E|L|O|S|P|E|D|C|I|O|L|K|K|L|R|L|P|N|I|P|H|Q|C|W|L|T|Q|Y|L|L|K|H|F| 248
OY 271 -----T|S|W|-----
Db 249 K|L|S|Q|A|S|K|N|L|N|A|V|L|S|E|I|F|S|P|V|L|F|F|P|A|S|S|D|N|T|E|H|L|I|K|A|I|E|I|L|I|S|T|E|N|E|R|O|P|A|P|A|L|P| 308
OY 274 -R|P|P|P|-----K|R|S|D|R|K|V|Q|H|N|E|W|I|G|E|Y|S|R|Q|A|V|E|A|F|M|K|E|N|K|G|S|F|L|R|D|C|S|T|K|S|K|E|E| 326
Db 309 P|K|P|K|E|T|V|A|N|N|S|N|N|N|S|L|Q|D|A|E|W|Y|W|G|D|S|R|E|E|V|N|E|K|-L|R|D|T|A|D|G|T|F|L|R|D|A|S|T|K|M|H|G|D| 367
OY 327 P|Y|V|L|A|V|E|Y|E|N|K|Y|Y|N|V|K|-I|R|F|L|E|R|N|O|F|A|L|G|T|G|L|R|G|D|E|K|F|D|S|Y|E|D|I|E|H|Y|K|N| 376
Db 368 -Y|T|L|T|---P|R|K|G|N|N|K|L|I|K|I|F|H|R|D|K|Y|G|F|S|D|P|L|---T|N|S|Y|E|L|I|N|H|Y|N| 410

```

Search completed: April 21, 2003, 12:37:57
 Job time : 29.3539 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:31 ; Search time 73.6821 Seconds

(Without alignments)
6314.347 Million cell updates/sec

Title: US-09-856-061-3

Perfect score: 2020

Sequence: 1 ttccagactcagctgcgc.....catcgactactacagaatt 1129

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09856061/runat_21042003_113019_3744/app_query.fasta_1.3150
-DB=SPREMBL_21 -OFMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -DOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09856061.qcgn.1.1_324.tunat_21042003_113019_3744 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEOUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertedrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	2020	100.0	376	Q9P2U9	Q9p2u9 homo sapien

ID	Q9P2U9	PRELIMINARY	PRT	376 AA	
AC	Q9P2U9				
DT	01-OCT-2000 (TREMBLrel. 15, Created)				
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	MIST (Fragment)				
CN	MIST				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20208989; PubMed=10744659;				
RA	Goltsuka R., Kanazashi H., Sasanuma H., Fujimura Y., Hidaka Y.,				
RA	Tatsuno A., Ra C., Hayashi K., Kitamura D.,				
RT	"A BASH/SLP-76-related adaptor protein MIST/Clunk involved in Ige				
RT	receptor-mediated mast cell degranulation."				
RI	Int. Immunol. 12:573-580(2000).				
DR	EMBL: AB032369; BAA96241.1; -				
DR	HSSP: P23727; 2PNB.				
DR	InterPro: IPR000980; SH2.				
DR	Pfam: PF00017; SH2; 1.				

ALIGNMENTS

RESULT 1

ID Q9P2U9 PRELIMINARY PRT: 376 AA.

Q9jnj3 mus musculu
Q9qez2 mus musculu
Q9ygc1 gallus gall
Q9d07 gallus gall
Q920L0 rattus norv
Q922M0 mus musculu
Q75499 homo sapien
Q75498 homo sapien
Q8wv28 homo sapien
Q9cun3 mus musculu
Q88504 mus musculu
Q9d413 mus musculu
Q54737 mus musculu
P70304 mus musculu
Q15054 homo sapien
Q8uuz2 xenopus lae
Q92809 abelion mur
Q91252 mus musculu
Q9n355 caenorhabdi
P70305 mus musculu
Q24708 drosophila
Q9cxc5 homo sapien
Q9utp6 schizosacch
Q9ypu1 drosophila
Q92529 homo sapien
Q8tarp2 homo sapien
Q95ym2 procambarus
Q9uph9 homo sapien
Q9ueg4 homo sapien
Q55032 mus musculu
P79956 xenopus lae
Q63789 rattus norv
Q8vny5 mus musculu
Q9tly2 canis famli
Q8r457 mus musculu
Q9ccl8 mus musculu
Q92279 mus musculu
Q99m51 mus musculu
Q9hcd7 homo sapien
Q92124 xenopus lae
Q64509 mus musculu
Q9cxc9 mus musculu
Q93j0 mus musculu
Q24218 drosophila

Db 199 GlnArgGlyProArgGlnArgSerAlaLysAspPheSerArgValLeuGlyAlaGluGlu 218
 Oy 595 GTTCCTCATACCGAGGAGGAGCCTGATCATCATCTCTTAGAAAACCAAACTACTCAA 654
 Db 219 GluSerHisHisIleThrLysProGluSerSerCysProSerSerSerAsnIleAsnThrGln 238
 Oy 655 GAGATTCCACTTGGCATTTAGCAGCTTCTTCATTCACGACAAACCAACGACGTCGCAAAAC 714
 Db 239 LysSerProProAlaIleAlaIleAserSerSerTyMetProGlyLysHisSerIleGlnAla 258
 Oy 715 AAGAGCATAGAGAGAGGAGCATGACGCCCTTCTCTCCAGAGATGCCAGCTCCAGCCAGC 774
 Db 259 ArgAspHisThrGlySerMetGlnHisCysProAlaGlnArgCysGlnAlaIleAlaIleSer 278
 Oy 775 TCGAGCCCTCAGCAAAATATACCTGCTTAATATACCAAGCTGGAGACCCTTTCCCC 834
 Db 279 HisSerPro-----ArgMetLeuProTyArgIleAsnThrAsnSerGlyLysProAspPro 296
 Oy 835 AAAAGTCTGATAGAAAGAGATGTCACGACAAATGAATGATGATGAGAAATACAGCCGC 894
 Db 297 ThrLysProAspGlyLysAspValTyrGlnAsnGlnTyrTyrIleGlyLysSerArg 316
 Oy 895 CAGGAGTGGAGAGGAGCATTCATGAAGGAGCAAGAGATGATGTTCTGTGCGAGAT 954
 Db 317 GlnAlaValGluAspValLeuMetLysGluAsnLysAspGlyThrPheLeuValArgAsp 336
 Oy 955 TGTTCACAAATCCAGAGAGAGCCCTATGTTTGGCTGTTTATGAGAACAAAGTC 1014
 Db 337 CysSerThrLysSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 356
 Oy 1015 TACATGTAAATCCGGCTCTCTGAGAGAGATTCAGCACTTGGCCCTGGGAGCAGACTC 1074
 Db 357 TyrAsnValLysIleArgPheLeuGluSerAsnGlnGlnPheAlaIleGlyThrGlyLeu 376
 Oy 1075 AGAGAGATGAGAACTTGTATCTACTAGAACATCATCATCACTAC 1122
 Db 377 ArgGlyAsnGluMetPheAspSerValGluAspIleIleGluHisTyr 392
 RESULT 4
 09YGC1 PRELIMINARY; PRT; 552 AA.
 AC 09YGC1;
 DT 01-MAY-1999 (TREMblrel. 10, created)
 DT 01-MAY-1999 (TREMblrel. 10, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE B cell linker protein BLNK.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99146381; PubMed=10023776;
 RA Ishii M., Kurosaki M., Pappu R., Okawa K., Ronko I., Fu C.,
 RA Shibata M., Iwamoto A., Chan A.C., Kurosaki T.,
 RT "BLNK required for coupling Syk to PLC gamma 2 and Rac1-JNK in B
 RT cells.";
 RL Immunity 10:117-125(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99049801; PubMed=9834055;
 RA Gotsuka R., Fujimura Y., Mamada H., Umeda A., Morimura T.,
 RA Uetsuka K., Doi K., Tsuji S., Kitemura D.,
 RT "BLNK, a novel signaling molecule preferentially expressed in B cells
 RT of the bursa of Fabricius.";
 RL J. Immunol. 161:5804-5808(1998).
 DR EMBL: AF089727; AAD12783.1;
 DR EMBL: AB015289; BA36275.1;
 DR HSSP: P08487; 2PLD.
 DR InterPro: IPR002965; P_rich_extensn.
 DR InterPro: IPR000980; SH2.

DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PRO1217; PRICHEXTENSN.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 SQ SEQUENCE 552 AA; 61822 MW; FB232179BE3BD072 CRC64;
 Alignment Scores:
 Pred. No.: 5,856-22 Length: 552
 Score: 334.00 Matches: 112
 Percent Similarity: 37.31% Conservative: 60
 Best Local Similarity: 24.30% Mismatches: 157
 Query Match: 16.53% Indels: 132
 DB: 13 Gaps: 15
 US-09-856-061-3 (1-1129) x 09YGC1 (1-552)
 Oy 94 TCGAAGAACTTGTGCTGAGTCTGATGAGCAAAAGCCACAT----- 141
 Db 83 TrpSerAspAspPheAspSerAspTyrGluAsnProAspGlyHisSerAspSerGluMet 102
 Oy 142 -----GATGATGACTATGATGACCTGAGCTTCGGATG 174
 Db 103 TyrValValProSerGluGluAsnProAspAspSerTyrGluProProProSerGluGln 122
 Oy 175 GAGAGCATGGCANTCATTAATTTTACCAGCCGCGCTATTAAGAAATCGAATAT 234
 Db 123 GluLysLys-----LysIleProSerSerPheProIleSerArgGlyLys 138
 Oy 235 GCAGAT-----ACACACTATTTCAGGTTCGAATGACACACCTCCCTCCGTTAGAC 285
 Db 139 AlaAspAsnArgThrSerHisGlnLeuProProIleAsnLysProLeuProSerThr 158
 Oy 286 ACCAGAGCTATAC-----TCCATTCAGACAGCCGACCTGGAACACA 327
 Db 159 ProSerSerAlaLeuProArgProLysLysProSerLeuProSerProAlaAlaLysPro 178
 Oy 328 CAGACGAGTTG-----GAAAGAGTGGCAAAACCAT 360
 Db 179 LysLeuProLeuLysProArgGluCysSerAspGluAspAsnTyrIleValProVal 198
 Oy 361 TCCAGGACGTCAGAGCCAAACATTAAAGAGATCGATCCGTA----- 405
 Db 199 AspAsnAspAspAspAsnTyrIleGluProThrGluSerSerThrProProAlaLys 218
 Oy 406 -----AGAAAGAACAGATTCCTTACCACTCTCT 435
 Db 219 ProProValAsnArgPheMetLysProProAlaLysSerAlaLeuProThrProProLys 238
 Oy 436 CGGCTCTCATMACACTTCCGMAAGATACCAACCTTGGCCCTGAGCCGAGAGCAGC 495
 Db 239 ProSerLeuAlaSerAspMetGlnGluValTyrGluValProGluGluGluGluLeu 258
 Oy 496 AGGCACCTTATATCCAGACACACCTTCCAGAGTCCAGGAATGCCAGT----- 549
 Db 259 SerProProProValThrArgPheThrLysPro-----LeuProAlaThrArg 274
 Oy 550 -----CAGATAAGCTTAAGGAGCTTAAGTACGATCTTGA 585
 Db 275 AlaGlnAsnAlaGluHisSerHisMetHisSerMetThrArgGluSerProLysLeuAsp 294
 Oy 586 GCAGAAAA-----GTTCTCATACAGCAGAGAGCCTGATCAACATCATCTG 633
 Db 295 AlserArgAsnIleLeuProLeuProArgAsnArgLeuHisProLysThrAspHisGlu 314
 Oy 634 TTGAAAC-----CAAAATAGCTCAAG----- 657
 Db 315 AlaAsnAsnAsnAspGluAsnHisSerPheSerAsnThrGlnGluSerLysPheProPro 334
 Oy 658 -----ATTCACTTGCATTAAGCATGTTCTCA----- 684
 Db 335 GlyAlaAlaProSerProLeuProAlaLeuLysLysThrSerAsnAlaValAsnPro 354

QY 685 -----TTCAAGACCAACACACATGTCGAAAC 714
 Db 355 AlalysProCysLeuProSerArgAspThrValLeuAspLysProThrAla 374
 QY 715 AGAGATCATAGAGAGCATGCAG-----CCCTGTCCTCCAGAGATGCCAGCCT 765
 Db 375 AlAspArgArgGlySerSerHisGluPheProLeuProLeuProSerGlyThr 394
 QY 766 CCAGCCAGC----- 774
 Db 395 ProLysSerSerLeuGlnLysProLeuValLeuProLysValProGluAlaProSerArg 414
 QY 775 -----TCAGCCCTCAGCAAAATATCTAGCCCTTAATATACACAGCTGCAGACCA 825
 Db 415 AlaleuGlyThrSerProHisSerIleSerIleSerSerThrAla----- 431
 QY 826 CCTTCCCAAGAGTGTGATAGAAGAGATGTCAGACACATGATGATGATGAGAA 885
 Db 432 -----AspGlnAspAlaGlyValHisSerLysAlaTyrIleAlaIleThr 446
 QY 886 TACAGCCCGCAGAGAGAGAGAGAGATGATGAGAGCAAGAGATGATGATGATG 945
 Db 447 CysAspArgLysThrAlaGluAspAlaLeuThrArgSerAsnLysAspGlySerPheLeu 466
 QY 946 GTCCGACATGTTCCACAAAATCCAGAGAGAGAGAGATGATGATGATGATGATG 1005
 Db 467 IleArgLysSerSerGlyGlnAspSerArgGlnProIleThrLeuValValPheThrAsn 486
 QY 1006 AACAAAGCTCAATGTAATAATCCGCTCTGAGAGAGAAATCAGACATGTCCTGGAG 1065
 Db 487 ArgArgValIleThrAsnIleProIleArgPheIleGluSerThrArgGlnIleThrAlaLeuGly 506
 QY 1066 ACAGAGACTCAGAGAGATGAGAAAGTTGATTCAGTAGAAGACATCATCAGACATCAG 1125
 Db 507 ArgGlyLysCysGlyGluGlnArgPheAspSerValAlaGluIleValGlnAsnHisGln 526
 QY 1126 AAT 1128
 Db 527 His 527

RESULT 5
 Q9DG07 PRELIMINARY: PRT: 530 AA.
 AC Q9DG07: 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE SLP-76 adaptor protein.
 GN SLP76.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC TISSUE=THYMUS.
 RX MEDLINE=20451093; PubMed=1093915;
 RA Ishai M., Kurosaki M., Inabe K., Chan A.C., Sugamura K., Kurosaki T.;
 RT "Involvement of LAT, Gads, and Grb2 in Compartmentation of SLP-76 to
 RL the Plasma Membrane.";
 DR J. Exp. Med. 192:847-856(2000).
 DR EMBL: AF226988; AAG18493.1; -;
 DR HSSP: P12931; ISHD.
 DR InterPro: IPR00160; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00454; SAM; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.

SO SEQUENCE 530 AA; 59580 MW; CF6CBCE8BD79909A CRC64;
 Alignment Scores:
 Pred. No.: 1.53e-20 Length: 530
 Score: 318.50 Matches: 107
 Percent Similarity: 37.59% Conservative: 55
 Best local Similarity: 24.83% Mismatches: 143
 Query Match: 15.77% Indels: 126
 Gaps: 15

US-09-856-061-3 (1-1129) x Q9DG07 (1-530)
 QY 120 GGATGAGCAAAAGGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 179
 Db 103 GlyTyrSerSerPheAspGlu-AspAspAspTyrGluSerProAspAspArgGlnGly 122
 QY 180 GACATGCGAGTCGATTAATAATTTACACGCCCGCCCTTAAAGAAATGTAATGACAGA 239
 Db 122 sGluAspGluAlaAspTyrGluSerProThrGluGluThrGluGluAlaGluHisAspSe 142
 QY 240 TACACACTATTTCAGAGTTGCAATGAGACATCCCTTCCTAGACACAGACCTCTAT 299
 Db 142 rAspGlyTyr-----GluProProProSerAsnAsnAspGluAlaHisH 157
 QY 300 CTCATGTCGACAGCCGACCTGGAACACACAGAGAGTGGAAAGAGTGGACAAACCCAT 359
 Db 157 sAsnValIlePheProAlaLysSerLeuAlaAsnAsnThrAspTyrIleAspArgPro-- 176
 QY 360 TTCAGGAGAGTCAGACCAACCAATTAAGAGATGATGATGATGATGATGATGATGATG 419
 Db 177 -----ProThrSerArgSerSerHisGlnPr 185
 QY 420 TCCTTACACCTCTCTCGGCC-----CT 443
 Db 185 oProValProProGlnAlaArgProGlyProSerProAlaProAlaSerPheGlyArgG 205
 QY 444 CATACACTCCGAGAGAGATGACACCTTG----- 474
 Db 205 yAlaSerLeuPro---AlaPheProProLeuProGlyAsnAsnAspArgAsnValLysPr 224
 QY 475 -----CCCCCTGAGCCGAGAGAGCAGC----- 498
 Db 224 oSerLysProProAlaProSerIleAspArgSerThrLysProProLeuAspArgLeuG 244
 QY 499 -CCACCTTATCTCAGAGACACACTTCCA---GAAGCCAGGAGATGCCAGCAGAT 554
 Db 244 yProProPheGluArgGluAsnProValProGlyArgLysProGlyHisProGlyLysLe 264
 QY 555 A-----AGCTTAAGAGAGCTTAAGTGAAGTCTTGAAGCAGAAAGTCTCATAA 605
 Db 264 uLeuThrProGlnLeuArgAlaLeuGlyGluGlnLeu-----AlaMetMetProLysPr 282
 QY 606 CCAGAGGAGAGCTGAATCAATCATCTGTTAGAAAACCAAAAT-----ACTCAAGAGAT 659
 Db 282 oProValProProSerAspArgTyrGluArgLysAsnProSerProLeuArgLysGlnI 302
 QY 660 TCCACTTGCCATTACAGAGTTTCATTCACGACCAAGCAACACAGTGCATGAACAGAGA 719
 Db 302 eProVal-----LysGlnGlyTyrPalaGlnGlnLysArgProGluGluGluLys 319
 QY 720 TCATAGAGAGAGCAGCAGCCCTGTTCTCTCAGAGATCCAGCTCCAGCAGCTGCGAG 779
 Db 319 pHis-----IleProGlnArgAlaValProGlnIleSerLeuPr 332
 QY 780 CCCTCAG---GAATAATATCTGCCCTATAAA----- 807
 Db 332 oProTyrSerSerAsnThrPheProSerLysSerIleLysAlaProProLysProGlySe 352
 QY 807 ----- 807
 Db 352 rAsnSerIleProGlyAlaGluSerAlaArgSerLeuSerAlaSerGlySerLeuProPr 372
 QY 808 -----TACCAAGCTGAGAC 824


```

OY 1033 TTCCTGAGAGGAATCAGACTTTCCTGGGACAGAGCACTCAGAGATGAGAGTTT 1092
DB 468 TyrgIngluGlusSerGlnValTyrlLeuLeuGlyThrGlyLeuArgGlyLysGlnsPhe 487
OY 1093 GATTCAGTAGAAGACATCATCGAACACTACAAG 1125
DB 488 LeuSerValSerAspIleLeuSerPyrPheArg 498

RESULT 7
O922M0 PRELIMINARY: PRT; 533 AA.
AC O922M0:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Lymphocyte cytosolic protein 2.
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC006948; AH06948.1;
DR MGD; MGI:1321402; LCP2.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 533 AA; 60238 MW; 50AEAD25EF0AD01 CRC64;

Alignment Scores:
Pred. No.: 3.58e-20 Length: 533
Score: 314.50 Matches: 121
Percent Similarity: 40.94% Conservative: 44
Best Local Similarity: 30.02% Mismatches: 169
Query Match: 15.57% Indels: 69
DB: 11 Gaps: 18

US-09-856-061-3 (1-1129) x O922M0 (1-533)
OY 58 GGCCAGTACCAAGAGATGAGCAAGCCTCTTGTAGCTGGAGAAAGACTTGTGTCGATC 117
DB 126 GlyAspTyrGlnSerProAsn-----GlnGlnGlnGlnAlaLeuVal 139
OY 118 CTGGATGGAGCAAAAGGCCACAGTGATGACTATGATGAGCCCTTCGATGAA 177
DB 140 AspAspAlaAla-----AspTyrGlnProProProSerAsnAsnGln 153
OY 178 GAGACATGGCAGTCGATTAATAATTTACAGCCCGGCTATA-----AAGAAATGAA 231
DB 154 GlnAlaLeuLysSer--SerIleLeuProProAsnSerPheHisAsnThrAsnSerMet 172
OY 232 TATGCAGAT---ACACACTATTTCAGAGTTCAGATGAGCACTCCCTT---CCGTTAGAC 285
DB 173 TyrIleAspArgProProThrGlyLysValSerGlnGlnProProValProProLeuArg 192
OY 286 ACCAGGACCTCTATCTCCATTGGACAGCCAGCTCGAACAACACAGACAGGTTGAAAGA 345
DB 193 ProLysProAlaLeuPro--ProLeuProThrGlyArgAsnHisSerProLeuSerPro 211
OY 346 GTGAGCAAAACCCATTTCCAGGAGCGTCAGAGCAAAACCTTAAAGAGATGATCCGTA 405
DB 212 ProHisProAsnHisGlnGlnProSerArgSerGlyAsnAsnLys----- 226
OY 406 AGAAAGACAAGATTCCTTTACCA-----CCTCCTCGGCTCTC 444
DB 227 ---ThrAlaLysLeuProAlaProSerIleAspArgSerThrLysProProLeuAspArg 245
OY 445 ATAACTTCCGGAAGAGTAGTACCAACCTTG-----CCCCCTGAGCCGGAG 489

```

```

DB 246 SerIleAlaProLeuAspArgGlnProPheIleLeuGlyLysLysProProPheSerAsp 265
OY 430 AGCAGCAGGCCACCTTTATCTCAGAGACACACCTTTCCAGAGAGTCCAG-----GGATG 543
DB 266 LysProSerAlaProLeuGlyArgGlnHis--LeuProLysIleGlnLysProProLeu 284
OY 544 CCCAGTCAGATAGCTTAAAGGACTTAAGTGAGCTT-----GAGCA 588
DB 285 ProProAlaMetAspArgHisGlnArgAsnGlnArgLeuGlyProValThrThrArgLys 304
OY 589 GAAAAGTTCCTCATACACAGAGAAAGCCTGATCACTCATCTGTAGAA----- 639
DB 305 ProProValProAlaGlnHisGlyArgGlyProAspArgArgGlnAsnAspArgVal 324
OY 640 AACCAAAATACTCAAGACATTCGCACTTGCATTTACAGTTCTTCTTACGACAGCAAC 699
DB 325 HisGlnArgProLeuProGlnProSerLeuProSerMetSerSerAsnThrPheProSer 344
OY 700 CACAGTGTCA----- 711
DB 345 ArgSerValGlnProSerSerLysAsnThrPheProLeuAlaHisMetProGlyAlaPhe 364
OY 712 AACAGAGATCATAGAGAGGAGCATGCAAGCCCTGTCT-----CTCAGAGATGCCAGCT 765
DB 365 SerGlnSerAsnIleGlyPheGlnGlnSerAlaSerLeuProProTyrPheSerGlnGly 384
OY 766 CCAGCAGCTGCAAGCCT-----CACGAAAATATCTGCCCTATTAATACACAGC 816
DB 385 ProGlyAsnArgProProLeuArgSerGlnGlyArgAsnLeuProLeuProValProAsn 404
OY 817 TGGAGACACCACTTTCACCAAGAGTGTGATAGAGATGTCACACAAATGAATGTTAC 876
DB 405 ---ArgProGlnProProSerProGlyGlnGlnGlnThrProLeuAspGlnGlnTyr 423
OY 877 ATTGGAATPACAGCCCGCAGCAGTGGAGAGCATTTATAGAGAAACAAGATGTT 936
DB 424 ValSerTyrIleThrArgProGlnAlaGlnAlaLeuArgLysIleAsnGlnsPglY 443
OY 937 AGTTCTTGTCGAGATGTTCACAAATTCACAGAGAGAGCCATATTTTGGCTGTG 996
DB 444 ThrPheLeuValArgAspSerSerLysIleThrAlaAsnAsnProTyrValLeuMetVal 463
OY 997 TTTTATGAGAAACATCTACATGTAAATCCGCTTCGTGAGAGATCAGCATTT 1056
DB 464 LeuTyrLysAspLysValTyrAsnIleGlnIleArgTyrGlnGlnGlnSerGlnValTyr 483
OY 1057 GCCCTGGGACAGAGATCAGAGAGATGAGAGATTGATTCAGTACAAACATCATCGAA 1116
DB 484 LeuLeuGlyThrGlyLeuArgGlyLysGlnAspPheLeuSerValSerAspIleLeuAsp 503
OY 1117 CACTACAAG 1125
DB 504 TyrPheArg 506

RESULT 8
O75499 PRELIMINARY: PRT; 433 AA.
ID O75499:
AC O75499:
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE B cell linker protein BLNK-S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation.

```

RL J. Biol. Chem. 272:27362-27368(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98361304; PubMed=9697839;
 RA Fu C., Turk C.W., Kurosaki T., Chan A.C.;
 RT "BLNK: a central linker protein in B cell activation."
 RL Immunity 9:93-103(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20050956; PubMed=10583958;
 RA Minegishi Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
 RT "An essential role for BLNK in human B cell development."
 RL Science 286:1954-1957(1999).
 DR EMBL: AF068181; AAC39937.1; -;
 DR EMBL: AF180756; AAF20383.1; -;
 DR EMBL: AF180740; AAF20383.1; JOINED.
 DR EMBL: AF180741; AAF20383.1; JOINED.
 DR EMBL: AF180742; AAF20383.1; JOINED.
 DR EMBL: AF180743; AAF20383.1; JOINED.
 DR EMBL: AF180744; AAF20383.1; JOINED.
 DR EMBL: AF180745; AAF20383.1; JOINED.
 DR EMBL: AF180746; AAF20383.1; JOINED.
 DR EMBL: AF180748; AAF20383.1; JOINED.
 DR EMBL: AF180749; AAF20383.1; JOINED.
 DR EMBL: AF180750; AAF20383.1; JOINED.
 DR EMBL: AF180751; AAF20383.1; JOINED.
 DR EMBL: AF180752; AAF20383.1; JOINED.
 DR EMBL: AF180753; AAF20383.1; JOINED.
 DR EMBL: AF180754; AAF20383.1; JOINED.
 DR EMBL: AF180755; AAF20383.1; JOINED.
 DR HSP: P29354; 18MB.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PFO0017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PSS0001; SH2; 1.
 SQ SEQUENCE 433 AA; 48229 MW; 0B36FE9FCF5DC7DC CRC64;
 Alignment Scores:
 Pred. No.: 4,9e-18 Length: 433
 Score: 291.00 Matches: 99
 Percent Similarity: 39.51% Conservative: 63
 Best Local Similarity: 24.15% Mismatches: 137
 Query Match: 14.41% Indels: 111
 DB: 4 Gaps: 16
 US-09-856-061-3 (1-1129) x 075499 (1-433)
 QY 60 CCAAGTACAGAGATGAACAGCCTCTTACAGCTGGGAAAGAACTTT----- 108
 DB 57 ProlaAspLugInglugIntrPserAsp-AspPheAspSerAspYrGlunProAs 76
 QY 109 -----GCTCAGTCTCGATGGAGCAAAAGCGACGATGATGACTA 132
 DB 76 PGLuHISerAspSerGluMetYrValMetProLaGluGluAsnAlaAspSperTy 96
 QY 153 TGAATGACCTGAGCTTCGATGGAAGACAGACAGTGCATTAAATTTTACACGCC 212
 DB 96 rGluProProProValGluInglu-----ThrArgProValHISProAla 112
 QY 213 GCCTATAAGAGATCTGAATATGACATACACATATTTCAGTTGCATGACACTCC 272
 DB 112 uProPheAlaArgGlyGluYrTleAspAsnArg-----SerSerGluArgH 128
 QY 273 CCTTCGTTAGACACAGAGACTCTATCTCATTTGACAGCGACCTGGAACACAGAC 332
 DB 128 sSerProProPheSerTyrThrLeuProSer-----LysProSerTyr----- 142
 QY 333 GAGCTTGGAAAGAGTGACAAACCCATTTCAGAGGACGTGAGAACCCAAATTAAGG 392
 DB 143 -----ProSerGluLysAlaArgLeuThrSerThrLeuProAl 155

QY 393 AGATGATCCGTAAAGAAAGAAAGATTCTTACCAGCCTCTGGGCTTCAPA----- 447
 DB 155 aleuThrAlaLeuGlnLysProGluValPro-----ProLysProLysGlyLeuLeuAs 174
 QY 448 -----ACACTCCAGAA 461
 DB 174 PGLuAlaAspTyrValValProValGluAspAsnAspGluAsnTyrTleHISProThrGl 194
 QY 462 GTACCAACCTTGGCCCTGAGCCG-----GAGAGCAGCAGGCC 500
 DB 194-UserSerSerProProGluLysGlyArgAsnSerGlyAlaTrpGluThrLysSerPr 214
 QY 501 ACCTTATCTCAGAGACACACCTTCCAGAGTCCAGGAATGCCAGTACG----- 552
 DB 214 oProProAlaAlaProSerProLeuProAlaGlyLysLysProThrThrProLeuLy 234
 QY 553 -----ATAAGTTAAGGACTTAAGTGAAGTCTTGAAGCAGAAAGTTCC 599
 DB 234 sThrThrProValAlaSerGlnGlnAsnAlaSerSerValCysGluGluLysProThrLe 254
 QY 600 TCATACACAGAGAGCCTAATCACTCTCTGTTAGAAACCAAAATCTCAGAGAT 659
 DB 254 o-----AlaGluArgHISArgLysSerSerHISArgGlnGlu----- 266
 QY 660 TCCACTTGCCATTAGCAGCTCTTCATTACAGACAGACACACAGTGTGCAAAACAGAGA 719
 DB 267 -----AlaValGlnSerProValPhe----- 273
 QY 720 TCATAGAGAGAGCAGCAGCCCTGTTCTCTCAGAGATGCCAGCTCCAGCAGTCCAG 779
 DB 274 -----ProProAlaGlnLysGl 279
 QY 780 CCCTCAGAAATATATAGTCCCTAT---AAATACACA---AGCTGAGACCA----- 825
 DB 279 nLleHISGlnLysProThrProLeuProAlaGlyPheHISGlnGlyLysProThrValAs 299
 QY 826 ---CCTTCCCAAAAGCTGTATGAAGATGTCAGCACAATGA----- 870
 DB 299 PGLyProLeuProSerPheSerSerAsnSerThrLISerGluGlnGluAlaGlyValLe 319
 QY 871 -----TGATACATTGAGATATACAGCCCGCAGGACGATGAGAGCATTCATGA 920
 DB 319 uCysLysProTrpTyrAlaGlyAlaCysAspArgLysSerAlaGluGluAlaLeuHISAr 339
 QY 921 GGAGAACAGAGATGATGTTGTCGAGATGTTGCCAAATCCAAAGAGAGGCC 980
 DB 339 gSerAsnLysAspGlySerPheLeuLeuArgLysSerSerGlnHISAspSerLysGlnPr 359
 QY 981 CTATGTTTGGCTGTGTTTATGAGAACAAAGTCTCAATGTAATAATCCGCTCTGGA 1040
 DB 359 oTyrThrLeuValValPhePheAsnLysArgValTyrAsnLleProValArgPheHISAr 379
 QY 1041 GAGCAATCAGCACTTGGCTGGGACAGAGACTCAGAGAGATGAGAGATTCATTCAGT 1100
 DB 379 uAlaThrLysGlnTyrAlaLeuGlyArgLysLysAsnGlyGluGlyTyrPheLysSerVa 399
 QY 1101 AGAAGACATCATGAACACTTACAAGAT 1128
 DB 399 lAlaGluIleIleArgAsnHISGlnHIS 408
 RESULT 9
 075498 PRELIMINARY; PRT; 456 AA.
 AC 075498;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE B cell linker protein BLNK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]

0 **AAGGCGCCCTGGTTCCTCCTCAAGA / 56**

DR HSSP; P16277; 1BLK.
MCD: WGT: 1018380; 40334040132015

RT Functional annotation of a full-length mouse cDNA collection.
 RL Nature 409:685-690(2001).
 DR EMBL: AK016886; BAB30480.1; -.
 DR HSSP: P16277; 1BLK.
 MD MGI:1918380; 4933424C13RLK.
 DR

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:14:41 ; Search time 77.294 Seconds
(without alignments)

5933.817 Million cell updates/sec

Title: US-09-856-061-1

Sequence: 1 acgagagcccaactgccccag.....taaaaaaaaaaaaaaaaa 1721

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer-n2p.model -DEV=alp
-Q=/cgn2_1/USPTO.spool/US09856061/rnuc.21042003_113018_3728/app_query.fasta.1.3150
-DB=A.geneseq.101002 -QMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPLC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosun62 -TRANS=human40.cd1
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MTLEN=0 -MAXLEN=2000000000
-USR=US09856061.ecgn.1.1.209.6runat.21042003_113018_3728 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6 -FCAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq.101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	75.8	435	22	Murine mast cell-s
2	1370.5	44.8	443	23	Human MIST protein
3	1316.5	43.1	428	23	Human MIST splice
4	1132	37.0	376	22	Human mast cell-sp
5	1109.5	36.3	353	23	Human MIST splice
6	326	10.7	456	22	Human protein sequ
7	319	10.4	449	19	Amino acid sequenc
8	319	10.4	456	19	Amino acid sequenc
9	287.5	9.4	457	19	Amino acid sequenc
10	260	8.5	95	20	Human SLF-76 SH2 d
11	213	7.0	84	21	Human secreted pro
12	212	6.9	46	22	Peptide #5655 enco
13	212	6.9	46	22	Protein #5353 enco
14	212	6.9	46	22	Human brain expres
15	212	6.9	46	22	Human bone marrow
16	212	6.9	46	22	Peptide #5407 enco
17	212	6.9	46	22	Peptide #5612 enco
18	212	6.9	46	23	Human peptide enco
19	195.5	6.4	503	22	Novel human diago
20	165.5	5.4	474	18	Novel human brain-specif
21	165.5	5.4	594	18	Human protein seq
22	150.5	4.9	743	22	Human protein seq
23	149.5	4.9	1353	22	Human protein seq
24	149.5	4.9	1357	22	Human protein seq
25	149.5	4.9	1356	22	Novel human diago
26	148.5	4.9	1396	22	Novel human diago
27	147	4.8	1408	22	Novel human diago
28	146.5	4.8	217	18	Novel human diago
29	146	4.8	1567	23	Novel human diago
30	146	4.8	1594	23	Novel human diago
31	144.5	4.7	474	18	Human ShcC mutant
32	144.5	4.7	474	18	Human ShcC mutant
33	142.5	4.7	474	18	Human ShcC mutant
34	142.5	4.7	474	18	Human ShcC mutant
35	141.5	4.6	474	18	Human ShcC mutant
36	141.5	4.6	474	18	Human ShcC mutant
37	141.5	4.6	474	18	Human ShcC mutant
38	141.5	4.6	474	18	Human ShcC mutant
39	141.5	4.6	474	18	Human ShcC mutant
40	141.5	4.6	474	18	Human ShcC mutant
41	141.5	4.6	474	18	Human ShcC mutant
42	141.5	4.6	474	18	Human ShcC mutant
43	141	4.6	728	22	Human ShcC mutant
44	140.5	4.6	469	17	Human ShcC mutant
45	140.5	4.6	474	18	Human ShcC mutant

ALIGNMENTS

RESULT 1

AAAB81070 standard; Protein: 435 AA.

AAAB81070;

25-JUN-2001 (first entry)

Murine mast cell-specific signal transduction protein.

Mast cell; signal transduction; mouse; allergic disease.

Mus musculus.

JP3146204-B1.

12-MAR-2001.

17-SEP-1999;

99JP-0263778.

XX 17-SEP-1999: 99JP-0263778.
 XX (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
 PA WPI: 2001-310022/23.
 DR N-PSDB: AAF86139.
 XX
 PT Mast cell-specific signal-transduction molecule, useful for screening
 PT therapeutic compounds for treating allergies, is specifically expressed
 PT by mouse mast cell.
 XX
 PS Claim 1: Page 7-8; 12pp: Japanese.
 XX
 CC This sequence represents a murine mast cell-specific signal transduction
 CC protein. The invention includes the cDNA and protein sequences of the
 CC mast cell-specific signal transduction molecule and an expression vector
 CC containing the polynucleotide sequence. The coding sequence of the signal
 CC transduction protein can be used for screening therapeutic compounds
 CC which will be useful for treating allergic diseases.
 XX

Sequence 435 AA:

Alignment Scores: 1.2e-203 Length: 435
 Pred. No.: 2316.00 Matches: 435
 Score: 2316.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 75.76% Gaps: 0
 DB: 22

US-09-856-061-1 (1-1721) x AAB81070 (1-435)

OY 255 ATGACACGACGAGGCAATATAAGACAAAGAAAGATGCGGTGATGATTCAG 314
 DB 1 MetThrSerGlnGlyAsnLysArgThrThrLysGluGlyPheGlyAspLeuArgPheGln 20
 OY 315 AACGTCCTCTGCTGAAAAATAGTCATGAGCCCAAGCCTCAGACAGTCCCAAGGGCGGTG 374
 DB 21 AsnValSerLeuLeuLysAsnArgSerTrpProSerLeuSerSerAlaLysGlyArgCys 40
 OY 375 CGAGCGGTCGTCGAAACACACTCCGATCAGACAGAAAGATTCGGGGGTCCTCCAGTGA 434
 DB 41 ArgAlaValLeuGlnProLeuProAspHisArgAsnLeuAlaGlyAlaProGlyGly 60
 OY 435 GAAAAATGCACACAGTACACAGCAGTACAGAAAGATCCTGATTCACCTGCTGAAGCATGG 494
 DB 61 GluLysCysAsnSerAsnAsnAspTyrGluAspProGluPheGlnLeuLeuLysAlaTrp 80
 OY 495 CCATCAATGAATAATTTTCCAGCCAGACCTATCCAGAAATGCGAATACGACATACAGC 554
 DB 81 ProSerMetLysIleLeuProAlaArgProIleGlnGlnSerGlnTyrAlaAspThrArg 100
 OY 555 TATTTCCAGATATGATGAGAGGCTCCCTTCGTTACCTCCCAAGGCTTCGTCGCACT 614
 DB 101 TyrPheGlnAspMetMetGlnAlaProLeuLeuLeuProProLysAlaSerValSerThr 120
 OY 615 GAGACACAAACAGGAGTGTAGATGACACAGCTGGAAGATGAGACAGCTTACTTTC 674
 DB 121 GlnAAGGlnThrArgAspValArgMetThrGlnLeuGlnGluValAspLysProThrPhe 140
 OY 675 AAGATGTCAGAAAGCAAGCGCTTAAAGATTCAAATACAAAAATAACAAAGACTCTTC 734
 DB 141 LysAspValArgSerGlnArgPheLysGlyPheLysTyrThrLysIleAsnLysThrPro 160
 OY 735 TTGCCACCTCCTCGGCTGCTATCCTACTCTCCCAAGAGATGACCAACCTTACCCCCACAGA 794
 DB 161 LeuProProProArgProAlaIleThrLeuProLysLysTyrGlnProLeuProProAla 180
 OY 795 CCACACAGAGGAGAGCAGTCTACTTGGCTCCAAAGCCACCTTCCAGAAAGTCCAGAG 854
 DB 181 ProProGlnGlnSerSerAlaTyrPheAlaProLysProThrPheProGluValGlnArg 200

OY 855 GGGCCAGGCGAGAGAGTGCAAAAAGACTTCAGTAGGGTCTTGAGACAGAAAGAAATCT 914
 DB 201 GlyProArgGlnArgSerAlaLysAspPheSerArgValLeuGlnGluGlnSer 220
 OY 915 CACCCACAGACAAAGCAGCAATCTTTTGGCCCTCATCAAAACCAACACACAGAAAGT 974
 DB 221 HisHisGlnThrLysProGlnSerSerCysProSerSerAsnGlnAsnThrGlnLysSer 240
 OY 975 CCACCTGCCATTGCCAGCTCTTCTTCATGACAGAAACACAGTATACAGCAGAGAC 1034
 DB 241 ProProAlaIleAlaSerSerSerTyrMetProGlyLysHisSerIleGlnAlaArgAsp 260
 OY 1035 CATACAGGTACATGACAGACTGTCTGCTCAGAGATGCCAAGCTGACAGCCACAGC 1094
 DB 261 HisThrGlySerMetGlnHisCysProAlaGlnArgCysGlnAlaAlaAlaSerHisSer 280
 OY 1095 CCTCGAATGCTCGCCCTATGAAAAACAACTGGGAAACCTGACCCCAAGCTCAT 1154
 DB 281 ProArgMetLeuProTyrGlnAsnThrAsnSerGluLysProAspProThrLysProAsp 300
 OY 1155 GAGAAAGATGCTGCGAGAAATGATGATGATGATGAGATGACAGTCCAGCAGCAGTGA 1214
 DB 301 GluLysAspValTrpGlnAsnGlnTyrTrpIleGlyGluTyrSerArgGlnAlaValGlu 320
 OY 1215 GATGCTTAAATGAAGAGACAAAGAGATGCTATTTTGGTCCGAGACTGCTTACAA 1274
 DB 321 AspValLeuMetLysGlnAsnLysAspGlyThrPheLeuValArgAspCysSerThrLys 340
 OY 1275 TCCAAAGCGCAACCATATGTTTGTGCTGCTTATGGAACAGCTGTACAAATGTGA 1334
 DB 341 SerLysAlaGlnProTyrValLeuValAlaPheTyrGlyAsnLysValTyrAsnValLys 360
 OY 1335 ATCCGTTTCTCTGAGACCAATACACAGTTCGCTGGCGACAGACTACAGAGAAATGAG 1394
 DB 361 IleArgPheLeuGlnSerAsnGlnInPheAlaLeuGlyThrGlyLeuArgGlyAsnGlu 380
 OY 1395 ATGTTGATTCGTGGAAGACATCATTAACACATCAATATTTCCATTCGTGATTA 1454
 DB 381 MetPheAspSerValGlnAspIleIleGlnHisTyrThrTyrPheProIleLeuLeuIle 400
 OY 1455 GATGGGAAAGACAGCGTCGACGACGACGAGAAACAGTCTACCTACAGCAGCTGCTC 1514
 DB 401 AspGlyLysAspLysAlaAlaArgArgLysGlnCysTyrLeuThrGlnProLeuProLeu 420
 OY 1515 GCCAGGCTCCTTCATCAGTACAGTCCAGCAGGACACTTCATGAG 1559
 DB 421 AlaArgLeuLeuThrGlnTyrSerSerGlnAlaLeuHisGlu 435

RESULT 2
 AAE22608
 ID AAE22608 standard; Protein: 443 AA.
 XX
 AC AAE22608;
 XX
 DT 26-JUL-2002 (first entry)
 XX
 DE Human MIST protein #1.
 XX
 KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
 KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cyostatic; immunomodulatory.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 84 /label= Tyrosine_phosphorylation_binding_site
 FT Modified-site 111 /label= Tyrosine_phosphorylation_binding_site
 FT Domain 306..311 /label= SH3_binding_proline-rich_motif
 FT Domain 324..407 /label= SH2_domain

```

XX  WO200226986-A2.
PM
XX
XX  04-APR-2002.
XX
XX  28-SEP-2001; 2001WO-US30593.
XX
XX  29-SEP-2000; 2000US-237030P.
XX
XX  (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX  Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
PI
XX  WPT. 2002-372126/40.
XX
XX  N-PSDB; AAD35800.
XX
XX  New isolated mast cell immunoreceptor signal transducer polypeptide,
PT  useful for treating immune disorder involving hyperactivity of B- or
PT  T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX
XX  Claim 11; Fig 3; 171pp; English.
XX
XX  The present invention relates to novel mast cell immunoreceptor signal
CC  transducer (MIST) proteins and polynucleotides encoding such proteins.
CC  MIST sequences of the invention are useful for preventing, treating or
CC  ameliorating a medical condition in mammalian subject. They are useful
CC  for treating an immune disorder involving hyperactivity of B- or T-
CC  lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
CC  lymphoma, tumor or thymoma in a mammal. MIST sequences are useful
CC  as targets for therapeutic intervention in immune cell disorders and
CC  inflammatory indications, for diagnosis and/or screening of disorders
CC  or diseases associated with expression of MIST, for screening for
CC  antagonists or inhibitors of the interaction of MIST with cellular
CC  signaling components. They are used in assays that detect activation
CC  or induction of various B and T-cell-related neoplasms or cancers.
CC  Sequences of the invention are also used in gene therapy. The present
CC  sequence is human MIST protein.
XX
XX  Sequence 443 AA:
SQ
XX
XX  Alignment Scores:
XX  Pred. No.: 7,59e-117 Length: 443
XX  Score: 1370.50 Matches: 278
XX  Percent Similarity: 73.30% Conservative: 46
XX  Best Local Similarity: 62.90% Mismatches: 113
XX  Query Match: 44.83% Indels: 5
XX  DB: 23 Gaps: 4
XX
XX  US-09-856-061-1 (1-1721) x AAE22608 (1-443)
XX
XX  213 GCTGAGTGAAGATCCCTCTTTCGTCAGTCCCAAGACATGACACAGGCGCAT 272
XX  |||||||
XX  Db 2 AAGGAGTGAAGATCCCTCTTTCGTCAGTCCCAAGACATGACACAGGCGCAT 21
XX  |||||||
XX  273 AAAAGCAGCAAGAAAGAGATTCGTCAGTCCCAAGACATGACACAGGCGCAT 332
XX  |||||||
XX  Db 22 ArglyshlrhrlysgllyserAsnAspLeuylsPheGlnAsnPheserLeuProlys 41
XX  |||||||
XX  333 AATAGTCATAGCCAGCCTCAGCAGTCCCAAGAGCGGCTGTCAGGCGTTTGGAAACA 392
XX  |||||||
XX  Db 42 AsnArgserTrpProArgIleAsnSerAlaThrGlyGlnTrgGlnArgMetAsnLysPro 61
XX  |||||||
XX  393 CTTCGGATCAGCAAGAACTTGTGGGTCCAGGTGAGAAATGCAATGCAAGTAAC 452
XX  |||||||
XX  Db 62 LeuLeuAspTrpGluArgAsnPhelaAlaValLeuAspGlyAlaLysGlyHisSerAsp 81
XX  |||||||
XX  453 AACCACTAGCAAGATCCCTGAGTCCAGTCCAGGACATGCAATGCAATGCAATTTTA 512
XX  |||||||
XX  Db 82 AspAspTrpArgAspAspProGluLeuArgMetGluGluTrpTrpGlnInsrtIleLysIleLeu 101
XX  |||||||
XX  513 CCAGCCAGACCTATTCAGGAATGCGAATAGCAGATACACGCTTTTTCACAGATATGATG 572
XX  |||||||
XX  Db 102 ProAlaArgProIleLysGluSerGluTrgAlaAspTrhHisTrgPheLysValAlaMet 121

```

```

OY 573 GAGGCTCCCTCTGTACTCTCCCAAGGCTTCTCTCTACCTGAGAGCAAAACAGGAT 632
XX  |||||
Db 122 AspTrhProLeuProLeuAspTrhArgTrhSerIleSerIleGlyGlnProTrhTrpAsn 141
XX  |||||
OY 633 GTGAGGATGACACAGCCTGGAAGAGTGACAGCCTTACCTTCAGAGATGCAAGCA 692
XX  |||||
Db 142 ThrGln---ThrArgLeuGlnArgValAspLysProIleSerLysAspValArgSerGln 160
XX  |||||
OY 693 CGCTTAAAGATTTCAATATACACAAATTAACAGACTCTTTCGCCACTCTCCGGCT 752
XX  |||||
Db 161 AsnIleLysGlyAspAlaSerValArgLysAsnLysIleProLeuProProTrpArgPro 180
XX  |||||
OY 753 GCTATCAGTCCGCCAAGAGATGACACCTTACCCCAAGCAGCAGAGAGAGCAGT 812
XX  |||||
Db 181 LeuIleTrhLeuProLysLysTrgGlnProLeuProPro---GluProGluSerArg 199
XX  |||||
OY 813 GCATACCTTCCGTCGCAAGGCCACTTTCGAGAGTCCAGAGGGGCCACAGCAGAGGT 872
XX  |||||
Db 200 ProProLeuSerGlnArgTrhTrhPheProGluValGlnArgMetProSerGlnIleSer 219
XX  |||||
OY 873 GCAAAAGACTTGAAGGTCTCTGAGCAGAGAAAGATCTCACCCAGCAGCAAGCA 932
XX  |||||
Db 220 LeuArgAspLeuSerGluValLeuGluAlaGluLysValProHisAsnGlnArgLysPro 239
XX  |||||
OY 933 GAATCTTCTGCCCATCATCAAAACCAACACAGAGAGTCCACTTCAGTCCAGC 992
XX  |||||
Db 240 GluSerTrhHisLeuLeuGlnAsnGlnAsnTrhGlnIleProLeuAlaIleSer 259
XX  |||||
OY 993 TCTTCTATCATGCGCAGAAAGACACTATATCAAGCAGACCATATACAGTATGAC 1052
XX  |||||
Db 260 SerSerPheTrhTrhSerAsnHisSerValGlnAsnArgAspHisArgLysIleMetGln 279
XX  |||||
OY 1053 CACTGCTCTGCTCAGAGATGCCAAGTGCAGCCAGCAGCAAGCCT-----CGAATGCTG 1106
XX  |||||
Db 280 ProCysSerProGlnArgGlySerGlnProAlaSerCysSerProHisGluAsnIleLeu 299
XX  |||||
OY 1107 CCTATGAAACACAACTCGGAGAAACCTGACCCCAAGAGCTGATGAAAGATGTC 1166
XX  |||||
Db 300 ProTrpLysTrpTrhSerTrpArgProProPheProLysArgSerAspArgVal 319
XX  |||||
OY 1167 TGGCAGATGATGATGATGATGATGAGATGACAGTCCAGGACAGTGAAGATGTTAATG 1226
XX  |||||
Db 320 GlnHisAsnGlnTrpTrpArgLysGlyLysArgGlnAlaValGluGlnAlaPheMet 339
XX  |||||
OY 1227 AAAGAGAACAGATGCTATTTTGGTCCGAGATGCTCTCAAAATCCAAAGCAGAA 1286
XX  |||||
Db 340 LysGluAsnLysAspLysSerPheLeuValArgAspCysSerTrhLysSerLysGluGln 359
XX  |||||
OY 1287 CCAATATGTTTGGTGGTGTATTATGGAACAGATGATGATGATGATGATGATGATG 1346
XX  |||||
Db 360 ProTrpValLeuAlaValAlaPheTrpGluAsnLysValTrpAsnValLysIleArgPheLeu 379
XX  |||||
OY 1347 GAGAGCAATGACAGTTCGCTGGGACAGACAGTCCAGAGAAATGATGATGATGATG 1406
XX  |||||
Db 380 GluArgAsnGlnGlnPheAlaLeuGlyTrhGlyLeuArgGlyLysArgLysPheAspSer 399
XX  |||||
OY 1407 GTGGAACATCAATGTAACATGACATATTTTCCATGTCGTAATAGATGGAAGAC 1466
XX  |||||
Db 400 ValGlnAspLysIleGlnHisTrpLysAsnPheProIleLeuIleAspGlyLysAsp 419
XX  |||||
OY 1467 AAG---GCTCAGCAGCAAGAAAGTGTCTACCTCAGCAGCAGCAGCAGTCTG 1523
XX  |||||
Db 420 LysTrhGlyValHisArgLysGlnCysHisLeuTrhGlnProLeuProLeuTrhArgHis 439
XX  |||||
OY 1524 CTTCCTC 1529
XX  |||||
Db 440 LeuLeu 441
XX
XX  RESULT 3
XX  AAE22609
XX  ID AAE22609 standard; Protein: 428 AA.
XX

```

AC	AAE22609;
XX	
DT	26-JUL-2002 (first entry)
XX	
DE	Human MIST splice variant protein from clone #7.
XX	
KW	Human; mast cell immunoreceptor signal transducer; MIST; immune disorder
XX	lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
KW	Inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX	
OS	Homo sapiens.
XX	
PN	WO200226986-A2.
XX	
PD	04-APR-2002.
XX	
PF	28-SEP-2001; 2001WO-0530593.
XX	
PR	29-SEP-2000; 2000US-237030P.
XX	
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
XX	
PI	Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX	
DR	WPI; 2002-372126/40.
XX	
DR	N-PSDB; AAD35801.
XX	
PT	New isolated mast cell immunoreceptor signal transducer polypeptide,
XX	useful for treating immune disorder involving hyperactivity of B- or
PT	T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX	
XX	Claim 11; Fig 5; 171pp; English.

The present invention relates to novel mast cell immunoreceptor signals transducer (MIST) proteins and polynucleotides encoding such proteins. MIST sequences of the invention are useful for preventing, treating or ameliorating a medical condition in mammalian subject. They are useful for treating an immune disorder involving hyperactivity of B- or T-lymphocytes in a mammal, for inhibiting growth of or inhibiting T-lymphoma, tumour or thymoma in a mammal. MIST sequences are useful as targets for therapeutic intervention in immune cell disorders and inflammatory indications, for diagnosis and/or screening of disorders or diseases associated with expression of MIST, for screening for antagonists or inhibitors of the interaction of MIST with cellular signalling components. They are used in assays that detect activation or induction of various B and T-cell-related neoplasms or cancers. Sequences of the invention are also used in gene therapy. The present sequence is human MIST splice variant protein from clone #7.

Sequence	428	AA:	
Alignment Scores:			
Pred. No.:	6,77e-112	Length:	428
Score:	1316.50	Matches:	266
Percent Similarity:	72.90%	Conservative:	46
Best Local Similarity:	62.15%	Mismatches:	111
Query Match:	43.07%	Indels:	5
Ds:	23	Gaps:	4

US-09-856-061-1 (1-1721) x AAE22609 (1-428)

QY 255 ATGACCAGCCAGGGGCATTAAGAAGCACACAGAAAGAGATTCCGTATCTGAGATTCCAG 314
||| |||||:::||||||| |||||:::|||||
Db 1 MetAsnArgGlnGlyAsnArgLysThrThrLysGluGlySerAsnAspLeuLysPheGln 20

```

315 AACGTCCTCTGCTGAAAAATATGGTCATAGGCCCAAGCCCTCAGCAGTGTCCCAAGGGGGCGGTGT 374
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
db 21 AsnPhseSerLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGlyGlnTyr 40

```

3/5 CGAGCGGTTTCGGACCACTTCGGATCACAGAAGAACTGGGCTGGGGTCCAGGTGGA 434
 :: :: ::||| ||| ||||| ||| |||
 41 GlnArgMetAsnLysProLeuLeuAspTrrGluArgAsnPheAlaAlaValLeuAspGly 60
 . Db

```

QY      435  GAAAAATGCAACAGTATACACAGCTACGAGATCTGTAGTTCACAGCTGTGTGAAGGCATGG 494
      |||  ::::::::::::::::::::::::::::  ::::  ::  |||
Db      61  AAlAysGlyHissSerAspAspArgTyrAspAspProGluLeuArgmetGluGluThrTrp 80

```

QY 495 CCATCATGAAATTTTACAGCCAGACCTATCCAGAAATCGGAATACGCAGATACAGC 554
||||:||||||||||||||||||||||||||||||||||||||||||
Db 81 GlnSerIleIysIleuProAlaIargProIleIysGlnSerGluTyrAlaIaspThrIls 1000

QY 555 TATTTCACGATATGANGGAGGCTCCCTTCTGTACCTCCCAAGGCTTCGTCCACT 614
 |||||::: ||||| || :: ||||:||||
 Db 101 TyrPhelysValAlamelAspThrProleuProleuAspThrArgThrSerIleSerIle 120

OY 615 GAGAGCAAAACCGAGGATGTAGAGATGACACACCTGGAAGAAGTGACAAAGCCTACCTTC 674
::: ||| ::: :: ||:::||||| |||||

675 AAGGATGTCAAGCCACGCTTTAAAGGATTCAAATACAAAATAAACAAGACTCCT 734

Db 140 LysaspValArgSerGlnAsnIleLysGlyaspPalaservaValArgLysasnLysIlePro 159
QY 735 TTGCACCTCTCGGCGCTGATACATCTCCCAAGAAAGTACCAACCTTACCCCGACGA 794

Db 160 LeuProProProArgProLeuIleThrLeuProLysLysTyrGlnProLeuProPro-- 178

```

179 GluProGluSerSerArgProProLeuSerGlnArgHisThrPheProGluValGlnArg 198
      |||||  |||  ::  ::  ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY      855  GGGCCAGCCAGAGAGCTGCAAAAGACTTCAGTAGGGCTCTTGAGCAGCAGAGAAGCATCT 914
        |||  |||  |||  ::|||  |||  |||||  |||||:::
Db      199  MetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnAlaGlnLysValPro 218

```

QY 915 CACCACGACAAAGCCAGATCTTCTTGCCCATCATCAACCACCAACACACAGAGAGT 974
||||:|||| |
Db 219 HisasnglnarqlvsProglIsertThisfeuleugluAsnglnAspTprglnuile 238
||||:|||| |

975 CCACCTGCCATTGGCAGCTCTTCCCTACATGCGCAGGAAGCACAGTATACAGCCAGAGAC 103

DB 239 PROEUALALLESSerSerSerPheThrThrSerAsnHisSerValGlnAsnArgAsp 258

QY 1035 CATACAGGTAAGCATCAGACACTGTCTCTCTCAGAGAGATGCCMACGTGCAGCCAGCCACAGC 109

Db	259	HISMRGLYGLYMECLINPROCYSSERPROGINARGCYGLINPROPRALASERCYSSER	278
Qy	1095	CCCT-----CGAAGCTGCCCTATGAAAACACAAACTGGGAGAAACCTGACCCCACTAAAG	1144

Db 279 ProHisGluAsnIleLeuProTyrIleTyrThrSerTrpArgProProIlePolysarg 298

Db 299 SerAspArgLysAspRvaIcGlnHisasnGlutPrtYrTleGlycUtuTyrSerAcrGlnAla 318
 ||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1149 CCGATGGACAGCATCTCTCCGACAAATGCATGTTCAGCAATACAGTCCCAGGCCA 1200

Oy 1209 GTGGAGATGTGTTAATGAAGAAGACAAGATGTACTTTTTTGGTCGAGACTGCTC 1266
|||||::: |||||||||:::|||||
Db 319 ValIGluAlaIlePheMetLysGLuSnLysASPglVserPheLeuValArgAsnCVssSer 338

QY 1269 ACAAAATCAGGCAGAACCATATGCTTTTGGTGTGTTTATGGGACAAGGCTCTACAAT 1322

DO 339 TheLyssefLysgLuGIuPrototyralLeuhalavalPhetYcgluasnllysalTYrasn 358
 OY 1329 GTCAAATCCGTTTCTCGAGAGCAATCAACAGTTTGGCCCTGGGACACAGACATACGAGCA 1386

Db 359 ValysIleArgPheLeuGluArgGlnGlnIlePheAlaLeuGlyThrGlyLeuArgGly 378

```
Db      379 AspgIuLysPheaspSPeValGIuAspIleIleGluHisTyrLysasnPheProIleLeu 398
```

Dy 1449 CTATTAGATGGGAAGAACAAG---GCTGCACGCCAGAGAAACACAGTGTACTCTACCACCCAGCCA 1505
|||||
Db 399 LeuILEASPGILYLYSASPRYLSTHRCIYVAIHISARGLYSGINCYSNHSLEUITHGLINPRO 418
|||||

QY 1506 CTGCGTCTGCGCAAGGCTCCTTCTC 1529

```

DB 419 LeuProLeuThrArgHisLeu 426
|||||
RESULT 4
AAB81071
ID AAB81071 standard; Protein; 376 AA.
XX
XX AAB81071;
XX
XX 25-JUN-2001 (first entry)
XX
XX Human mast cell-specific immunoreceptor signal transducer.
XX
XX Mast cell; signal transduction; human; allergic disease; MIST;
XX mast cell-specific immunoreceptor signal transducer.
XX
XX Homo sapiens.
XX
XX JP3146204-B1.
XX
XX 12-MAR-2001.
XX
XX 17-SEP-1999; 99JP-0263778.
XX
XX 17-SEP-1999; 99JP-0263778.
XX
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
XX MPI: 2001-310022/33.
XX
XX N-PSDB; AAF86140.
XX
XX Mast cell-specific signal-transduction molecule, useful for screening
XX therapeutic compounds for treating allergies, is specifically expressed
XX by mouse mast cell.
XX
XX Example 1; Page 10-11; 12pp; Japanese.
XX
XX This sequence represents human MIST (mast cell-specific immunoreceptor
XX signal transducer). The invention relates to cDNA encoding a murine mast
XX cell-specific signal transduction protein. Included in the invention are
XX cDNA and protein sequences of the mast cell-specific signal transduction
XX molecule and an expression vector containing the polynucleotide sequence.
XX The coding sequence of the signal transduction protein can be used for
XX screening therapeutic compounds which will be useful for treating
XX allergic diseases.
XX
XX Sequence 376 AA;
XX
XX
XX Alignment Scores:
XX Pred. No.: 5,53e-95 Length: 376
XX Score: 1132.00 Matches: 229
XX Percent Similarity: 72.07% Conservative: 42
XX Best Local Similarity: 60.90% Mismatches: 101
XX Query Match: 37.03% Indels: 4
XX DB: 22 Gaps: 3
XX
XX US-09-856-061-1 (1-1721) x AAB81071 (1-376)
XX
XX 309 TTCGAGAACGTCCTCTCTCTGAAATATGTCATGCGCCACGCTCAGCAGTCCCAAGG 368
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 1 PheGlnAsnPheSerLeuProLysAsnArgSerTrpProAlaGlnAsnSerAlaThrIly 20
XX
XX 369 CGGTGCGAGGGGTCTTGGAGCACTTCGATCCAGACAGAACTTGGCTGGGGTCCCA 428
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 21 GlnTyrGlnArgMetAsnIlyProLeuLeuAspTrpGlnArgAsnPheAlaIleValLeu 40
XX
XX 429 GGTGAGAAAATGCAACAGTACAGACAGTACGAGATCTGAGTTCACGCTGCTGAG 488
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 41 AspGlyAlaIlyGlyHisSerAspAspIlyAspAspProGlnLeuArgMetIleuGln 60
XX
XX 489 GATGCGCATCAATGAAATTTTACGACGACACTATCCAGAGATCGGAATACGACAGT 548
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 ThrTrpGlnSerIleuLeuProAlaArgProIleLysGlnSerGlnTyrAlaAsp 80

```

```

QY 549 ACACGCTATTTCAGATATGATGAGAGTCCCTTCTGTATCTCCACGAGCTTCTGC 608
DB 81 ThrHisTyrPheIlyValAlaIleAspThrProLeuPheIlyAspThrArgHisSerIle 100
QY 609 TCCACTGAGAGACAAACGAGGATGTGAGATGACACAGCTTGAGAAAGTGACAGGCT 668
DB 101 SerIleGlyGlnProThrTrpAsnThrGln---ThrArgLeuGlnArgValAspLysPro 119
QY 669 ACCCTTCAGAGATGTCAGAAAGCCACGCTTTAAAGATTCMAATACCAAAATTAACAG 728
DB 120 IleSerArgAspValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLys 139
QY 729 ACTGCTTGGCCACCTCCCTGGGCTGCTATCAGTCCGCCCAAGATACCAACCTTACCC 788
DB 140 IleProLeuProProArgProLeuIleThrIleProLysIlyGlnProLeuPro 159
QY 789 CCAGCACACACAGAGAGAGAGAGTGCATCTTCCGCTCCAAAGCCACCTTCCAGAGTC 848
DB 160 Pro---GlnProGlnSerSerArgProProLeuSerGlnArgHisThrPheProGlnVal 178
QY 849 CAGAGGGGGCCAGCAGAGAGAGTGCAAAAGCTTACAGTACGCTCTTGACAGAGAA 908
DB 179 GlnGlyMetProSerGlnIleSerLeuArgAspLeuSerGlnValIleuGlnIleLys 198
QY 909 GAATCTCACCCACAGACAAAGCCAGATCTTCTTCCATCAACCAACCAACACACAG 968
DB 199 ValProHisAsnGlnArgLysProGlnSerThrHisLeuLeuGlnAsnIleAsnThrGln 218
QY 969 AAGAGTCCACCTCCATTTGCCAGCTCTCTTACATCCAGAAAGCAGATATACAGCC 1028
DB 219 GlnIleProLeuAlaIleSerSerSerPheThrThrSerAsnIleSerValGlnAsn 238
QY 1029 AGAGACATATCAGTACATGACAGACAGTCCGCTCAGAGAGCCAGCTGACAGCCAGC 1088
DB 239 ArgAspHisArgGlyGlyMetGlnProCysSerProGlnArgGlyGlnProProAlaSer 258
QY 1089 CACAGCCCT-----CGAATGCTGCCCTATGAAACAAACAACTGGAGAAACCTGACCC 1142
DB 259 CysSerProHisGlnAsnIleLeuProTyrLysThrSerTrpArgProProPhePro 278
QY 1143 ACAAGCCTATGAGAGAGATGTCTGCGACAAATGATGTACATTTGGAGAAATACACTCC 1202
DB 279 LysArgSerAspArgLysAspValGlnHisAsnGlnIlyPyrIleGlyGlnTyrSerArg 298
QY 1203 CAGGCACTGGAGATGCTTAATGAAGAGAACAGAGATGTAATTTTGTGCTCGAGAC 1262
DB 299 GlnIleValGlnGlnAlaPheMetLysGlnAsnLysAspLysSerPheLeuValArgAsp 318
QY 1263 TGCCTACAAAATCCAGAGCAGACCATATGTTTGGTGTGTTATGGGAACAGATC 1322
DB 319 CysSerThrLysSerLysGlnGlnProTyrValLeuAlaValPheTyrGlnAsnLysVal 338
QY 1323 TACAATGTGAAAATCCCTTCCCTGAGAGACAAACAGTATGGCCCTGGGACAGAGCTA 1382
DB 339 TyrAsnValLysIleArgPheLeuGlnArgAsnGlnIlePheAlaLeuGlyThrHisGlyLeu 358
QY 1383 CGAGGAATGAGATGTTGATTTCTGTGGAAGACATCATTTGAAACATAC 1430
DB 359 ArgGlyLysProlLysPheAspSerValGlnAspIleIleGlnHisTyr 374

```

```

RESULT 5
AAE22610
ID AAE22610 standard; Protein; 353 AA.
XX
XX AAE22610;
XX
XX 26-JUL-2002 (first entry)
XX
XX Human MIST splice variant protein from clone #12.
XX
XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
XX lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;

```

KM Inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
 XX Homo sapiens.
 OS
 XX MO200226986-A2.
 PN
 XX
 04-APR-2002.
 PD
 XX
 28-SEP-2001; 2001WO-US30593.
 PE
 XX
 29-SEP-2000; 2000US-237030P.
 PR
 XX
 (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA
 PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
 XX
 DR WPI; 2002-372126/40.
 DR N-PSDB: AAD35802.
 XX
 PT New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
 PS
 XX Claim 11; Fig 8; 171pp; English.

The present invention relates to novel mast cell immunoreceptor signal
 transducer (MIST) proteins and polynucleotides encoding such proteins.
 MIST sequences of the invention are useful for preventing, treating or
 ameliorating a medical condition in mammalian subject. They are useful
 for treating an immune disorder involving hyperactivity of B- or T-
 lymphocytes in a mammal, for inhibiting growth of B- or T-
 lymphoma, tumor or thymoma in a mammal. MIST sequences are useful
 as targets for therapeutic intervention in immune cell disorders and
 inflammatory indications, for diagnosis and/or screening of disorders
 or diseases associated with expression of MIST, for screening for
 antagonists or inhibitors of the interaction of MIST with cellular
 signaling components. They are used in assays that detect activation
 or induction of various B and T-cell-related neoplasms or cancers.
 Sequences of the invention are also used in gene therapy. The present
 sequence is human MIST splice variant protein from clone #12.

XX Sequence 353 AA;

Alignment Scores:

Pred. No.: 6.25e-93 Length: 353
 Score: 1109.50 Matches: 226
 Percent Similarity: 73.22% Conservative: 31
 Best Local Similarity: 64.39% Mismatches: 89
 Query Match: 36.29% Indels: 5
 DB: 23 Gaps: 4

US-09-856-061-1 (1-1721) x AAE22610 (1-353)

QY 486 AAGGCATGCCATCAATGAAATTTTACAGCCAGACCTATTCAGATGGGATACGA 545
 Db 3 GlutThrtPrgInserIleIleLeuProIleArgProIleIleGluSerGluTyrAla 22
 QY 546 GATCAGCGCATTTCCAGGATATGATGAGGCTCCCTGTGTTACCTCCCAAGGCTTCT 605
 Db 23 AsphThrtIstYrPheIleValAlaMetCaspThrProLeuProLeuAspThrArgThSer 42
 QY 606 GTCTCCATGAGACAAACCCAGGATGATGACAGCTGACAGCGTGAAGAAGTGAGCAAG 665
 Db 43 IleSerIleGluInProThrtPheInThrGln---ThrtGluGluAlaGValAspLys 61
 QY 666 CTRACCTTCAAGATGTGCAAGCAAGCAAGCTTTAAAGATTCAATACACAAATTAAC 725
 Db 62 ProIleSerIleYsAspValArgSerGlnAsnIleIleGluIleAspValArgLysAsn 81
 QY 726 AAGACTCCTTTGGACACTCTCGGCGCTGATACACTCTCCCAAGAAGTACCAACCTTA 785
 Db 82 LysIleIleProIleuProIleuArgProIleuIleThrLeuProLysIleGlnProLeu 101

QY 786 CCCCCAGCACACAGAGAGAGACAGTGCATCTTCGCTCCAAAGCCACCTTTCAGAA 845
 Db 102 ProPro---GluProGluSerSerArgProProLeuSerGlnArgHisThrPheProGlu 120
 QY 846 GTCCAGAGGGGGCCAGGAGAGAGAGTGCAGAAAGACTTCAGTAGGCTCTTGACACAGA 905
 Db 121 ValGlnArgMetProSerGlnIleSerIleuArgAspLeuSerGluValLeuGluAlaGlu 140
 QY 906 GAAGAATCTCACACACAGCAAGCAAGCAAGATCTTCTGCCATCATCAACCAACCAACA 965
 Db 141 LysValProHisAsnGlnArgLysProGluSerThrHisLeuGlnAsnGlnAsnThr 160
 QY 966 CAGAGAGTCCACCTGCGCATTCGACGCTCTTCTTCATGATCCAGCAAGCAAGTATACAA 1025
 Db 161 GlnGluIleProLeuAlaIleSerSerSerPheThrIleSerAsnHisSerValGln 180
 QY 1026 GCCAGAGACATACAGATAGCATGACAGACTGCTGCTGCTGAGATGCCAAGCTGCACCC 1085
 Db 181 AsnArgAspHisArgGluGlyMetGlnProCysSerProGlnArgCysGlnProProAla 200
 QY 1086 AGCCACAGCCCT-----CGAATGCTGCCCTATGAAACACAAACCTCGAAGAACCTGAC 1139
 Db 201 SerCysSerProHisGluAsnIleLeuProTyrLysTyrThrSerTrpArgProProPhe 220
 QY 1140 CCCACAAAGCCTGATGAGAGAGATGCTGCGCAGATGATGATGATGATGATGATGATGAT 1199
 Db 221 ProLysArgSerAspArgLysAspValGlnHisAsnGluTyrLysGlyIleGlyTyrSer 240
 QY 1200 CGCCAGCGCATGTGAAGATGTGTTATGAGAAAGCAAGATGTGTTATGTTGTCGCGA 1259
 Db 241 ArgGlnAlaValGlnGluAlaIlePheMetLysGluAsnLysAspGlySerPheLeuValArg 260
 QY 1260 GACGCTCTTCAAAATTCAGAGCAAGCAAGCATATGTTGTTGTTGTTATGAGAACAG 1319
 Db 261 AspYsserThrLysSerLysGluGluProTyrValLeuAlaValPheTyrGluAsnLys 280
 QY 1320 GTCTACATGTGAATAATCCGTTCTCGAGAGCATACAGATGTTGCGGGGACAGCA 1379
 Db 281 ValIleAsnValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlyThrGly 300
 QY 1380 CTACGAGAAATGATGATGTTGATCTGTGGAACATCATTCAGACATATATTTT 1439
 Db 301 LeuArgLysAspGluLysPheAspSerValGlnAspIleIleGlnHisTyrLysAsnPhe 320
 QY 1440 CCCATCTGCAATAGATGGGAAAGACAAAG---GCTGCAGCGCAAGAAAGCTCTACCTC 1496
 Db 321 ProIleIleLeuIleAspGlyLysAspLysThrGlyValHisArgLysGlnCysHisLeu 340
 QY 1497 ACCCAGCCACTGCTCTGCGCAGGCTCCTTCTC 1529
 Db 341 ThrGlnProLeuProLeuThrArgHisLeuLeu 351

RESULT 6
 AAB93406
 ID AAB93406 standard; Protein; 456 AA.
 XX
 AC AAB93406;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:12602.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 PS Claim 8: SEQ ID 12602; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH9893 represent human amino acid sequences; and AAH13629 to AAH1632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 XX Sequence 456 AA:
 SQ
 Alignment Scores:
 Pred. No.: 5 91e-21 Length: 456
 Score: 326.00 Matches: 99
 Percent Similarity: 37.06% Conservative: 50
 Best Local Similarity: 24.63% Mismatches: 141
 Query Match: 10.66% Indels: 112
 DB: 22 Gaps: 9
 US-09-856-061-1 (1-1721) x AAB93406 (1-456)
 QY 444 AACAGTACACGACGACTAGAGAGCTGAGTCCAGCTGCGAAGCGATGCCATCATG 503
 DB 91 AsnAlaAspSerPheTyrGluProProProValGlu-----GlnGluThrArg 106
 QY 504 AAAATTTTACGACGACGACGATCCAGATGCGAATACGACGATATTTCCAG 563
 DB 107 ProValHisProAlaLeuProPheAlaArgGlyGluTyrIleAspAsnArgSerSerGln 126
 QY 564 GATATGATGAGAGGCTCC-----CTTCTGTTACCTCCAGAGCTTGTCTCCACT 614
 DB 127 -----ArgHisSerProProPheSerIleThrLeuProSerIleProSer 144
 QY 615 GAGAGACAAACAGGATGTGAGATGACACAGCTGGAAGAAGTGACACCTACCTTC 674
 DB 145 GluLysAlaArgLeuThrSerThrLeuProAlaLeuThrAlaLeuGlnLysProGlnVal 164
 QY 675 AAGGATGTACAGACCAACGCTTTAAAGATTCAAATACAAAAATAAACAAGACTGCT 734
 DB 164 ----- 164

QY 735 TTGCACCTCCGCGCCCTGCTATCATCTCCCAAGAGTACCAACCTTACCCACAGA 794
 DB 165 -----ProProLysProLysGlyLeuLeuGlnLysGlnAlaAspTyrValValProVal 182
 QY 795 CCACGACGAGAGAGGAGTGCATCTTCCCTCCAAAGCCACCTTTCAGAAATCCAGAG 854
 DB 183 GluAspAsnAspGlnAsnTyrIleHisProThrGlnSerSerProProGlnLys 202
 QY 855 GGGCCCGAGGACAGAGATGTCAAAGACTTCAGTAGGCTCTTGGACGAAAGAAATCT 914
 DB 203 AlaProMet-----Val 206
 QY 915 CACCCACGACAAAGCCAGAAATCTTGTCCCATCATCAAAACCAACACA----- 965
 DB 207 AsnArgSerThrLysProLysSerSerThrProAlaSerProProGlyThrAlaSerGly 226
 QY 966 -----CAGAAGAGTCCAGCTGCCATTCAGCTCT----- 995
 DB 227 ArgAsnSerGlyAlaTyrPglThrLysSerProProAlaAlaProSerProLeuPro 246
 QY 995 ----- 995
 DB 247 ArgAlaGlyLysLysProThrThrProLeuLysThrThrProValAlaSerGlnAsn 266
 QY 996 -----TCTACATGCGAAGAAAGACAGTATACAAAGCAGACCATACAGTACAGTCGAG 1052
 DB 267 AlaSerSerValCysGlyLysLysProLysProLysProLysGlnAlaArgHisArgLysSerHis 286
 QY 1053 CACTGCTGCTCTCAGAGATCCCAAGCTGCAGCCAGCCAGCCCTGGAATGCTGCCCTAT 1112
 DB 287 -----ArgGlnGlnAlaValGlnSerProValPheProAla 299
 QY 1113 GAAACACAAACTCGGAGAACTGACGCC----- 1142
 DB 300 GlnLysGlnIleHisGlnLysProLysProLysProLysProLysGlnLysPro 319
 QY 1143 -----ACAAAGCTGATGAGANG 1160
 DB 320 ThrValAspLysProLeuProSerPheSerSerAsnSerThrIleSerGlnGlnAla 339
 QY 1161 GATGCTGCGACGAATGATGCTATGCGAATACAGTACAGTCCAGCTGAGAGTGC 1220
 DB 340 GlyValLeuCysLysProTyrPheAlaGlyAlaCysAspArgLysSerAlaGlnGlnAla 359
 QY 1221 TTAATGAAGAGACAGAGATGCTACTTTTGTGCTCCGAGACTGCTCAAAATCCAG 1280
 DB 360 LeuHisArgSerAsnLysAspLysSerPheLeuIleArgLysSerSerGlyHisAspSer 379
 QY 1281 GCAGAACCATATGTTGTGTGTGTTTATGGAACAGGCTTACATGCTGAATCCGT 1340
 DB 380 LysGlnProTyrThrLeuValAlaPhePheAsnLysArgValLysIleProValArg 399
 QY 1341 TTCCTGAGAGCAATCAACAGTTTGGCTGGCCGACAGAGACTAGAGAAATGATGTT 1400
 DB 400 PheIleGlnAlaThrLysGlnTyrAlaLeuGlnArgLysLysAsnGlyGlnGluTyrPhe 419
 QY 1401 GATTCGTGGAAGACATCATGAACTACATCACTATATTTCCATTCGTGTAATAGATGG 1460
 DB 420 GlySerValAlaGluIleIleArgAsnHisGlnHisSerProLeuValLeuIleAspSer 439
 QY 1461 AAAGAC 1466
 DB 440 GlnAsn 441
 RESULT 7
 AAM59866 standard; Protein: 449 AA.
 AC AAM59866;
 AC AAM59866;
 DT 20-NOV-1998 (first entry)

KW ras pathway; GDP; GTP; calcium pathway; antigen.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 XX FT Misc-difference 327
 FT /note= "encoded by AGC"
 XX
 XX MO9832852-A1.
 XX
 XX 30-JUL-1998.
 XX
 XX 23-JAN-1998; 98MO-US01394.
 XX
 XX 17-MAR-1997; 97US-0819013.
 XX 24-JAN-1997; 97US-0788322.
 XX
 XX (UNIW) UNIV WASHINGTON.
 XX
 XX Chan AC, Fu C;
 XX
 XX WPI; 1998-427948/36.
 XX
 XX N-PSDB; AAV41901.
 XX
 XX Human B cell linker proteins - useful in the treatment of diseases
 XX involving increased or decreased apoptosis
 XX
 XX Claim 4; Fig 1; 56pp; English.
 XX
 XX This is the amino acid sequence of the human BLNK-1 (B cells linker
 XX protein-1) protein, used the treatment of diseases involving the
 XX increase and decrease of apoptosis, in the method of the invention.
 XX BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
 XX intracellular calcium levels and Grb2 is an adapter molecule
 XX containing two SH3 domains that mediate its interaction with the
 XX guanine nucleotide exchange factor, Son of Sevenless (SOS) which in
 XX turn activates the ras pathway by facilitating the exchange of GDP for
 XX GTP on the ras molecule. Activation of both ras and calcium pathways
 XX are required for efficient B cell antigen receptor function. BLNK
 XX binding proteins can be used to identify BLNK proteins in a target
 XX sample.
 XX
 XX Sequence 456 AA;
 XX
 XX Alignment Scores:
 XX Pred. No.: 2.6e-20 Length: 456
 XX Score: 319.00 Matches: 99
 XX Percent Similarity: 39.65% Conservative: 58
 XX Best Local Similarity: 25.00% Mismatches: 139
 XX Query Match: 10.44% Indels: 100
 XX DB: 19 Gaps: 12
 XX
 XX US-09-856-061-1 (1-1721) x AAM59865 (1-456)
 OY 444 AACAGTAACAGAGACTAGTCTGAGTCCAGCTGTCGATGAGCATGCCATCAATG 503
 DB 91 AAAAAAAAAAASPSerTyrGluProProValGlu-----GlnGluThrArg 106
 OY 504 AAAATTTTACCAAGCAGACTATCCAGAAATGGATATACAGATACAGCTATTTCAG 563
 DB 107 ProValHisProAlaLeuProPheAlaArgGluTyrIleAspAsnArgSerGln 126
 OY 564 GATATGATGAGGCTCC-----CTTCTGTACCTCCCAAGGCTTCTGTCTCCAT 614
 DB 127 -----ArgHisSerProProPheSerLysThrLeuProSerLysProSerTrpProSer 144
 OY 615 GAGAGCAAAACAGAGATGAGATGATACACAGCTGAGAAAGTGGACAGCTTACCTTC 674
 DB 145 GluLysAlaIleArgLeuThrSerThrLeuProAlaLeuThrAlaLeuGlnLysProGlnVal 164
 OY 675 AAGGATGTCTGAAAGCCAGCGCTTAAAGATTCAAAATACACAAAATAAACAAGACTCCT 734
 DB 164 ----- 164

OY 735 TTGGACACCTCCGGGCTGCTATCAGCTTCGCCAGAGATACCAACCTTACCCGAGA 794
 DB 165 -----ProProLysProLysGlyLeuLeuGluAspGlnAlaAspTyrValValProVal 182
 OY 795 CCACAGAGAGAGAGAGATGCTATCTTCGCTCCAAAGCCACCTTCCAGAAATCCAGAG 854
 DB 183 GluAspAsnAspGluAsnTyrIleHisProThrGluSerSerSerProProGlnLys 202
 OY 855 GGGCCCAAGCAGAGAGAGACTGCAGAAAAGACTTCAATGAGGCTCTGGAGCAGAGAAATCT 914
 DB 203 AlaProMet-----Val 206
 OY 915 CACCAACAGCAAAAGCAGAGATCTTTCGCCATCATCAAAACCAAAACACA----- 965
 DB 207 ArgAsnSerThrLysProAsnSerSerThrProAlaSerProProGlyThrAlaSerGly 226
 OY 966 -----CAGAGAGTCACCTCCATTCGACAGCTTCTTCATATG 1004
 DB 227 ArgAsnSerGlyAlaTyrGluThrLysSerProProPro-----AlaAlaProSerProLeu 245
 OY 1005 CCA-----GGAAGCAC-----AGTATACAAAGCCAGAGAC 1034
 DB 246 ProArgAlaGlyLysLysProThrThrProLeuLysThrThrProValAlaSerGlnGln 265
 OY 1035 CATACAGGTAGCATG-----CAGCACTGCTCCTGCTCAGATGCCAAGCTGCA--- 1082
 DB 266 AsnAlaSerSerValCysGluGluLysProLeuProAlaGluArgHisArgLysSerSer 285
 OY 1083 -----GCCAGCAGACGCCCTCGAATGCTGCCCTATGAAACAAACTCGGAG 1130
 DB 286 HisArgGlnGluAlaValGlnSerProValPheProProAlaGlnLysGlnIleHisGln 305
 OY 1131 AAACCTGACCCC----- 1142
 DB 306 LysProLysProLeuProAlaGlyPheThrGluGlyLysAsnProThrValAspGlyProLeu 325
 OY 1143 -----ACAAAGCTGATGAGAGATGCTGTGTCGACAGAAATGAA 1178
 DB 326 ProIlePheSerSerAsnSerThrIleSerGlnGlnGlnAlaGlyValLeuCysLysPro 345
 OY 1179 TGTACATTTGAGAGATACAGTCCGCCAGGCTGCAAGATGTTTAAATGAAGAGACAG 1238
 DB 346 TrpTyrAlaGlyAlaCysAspArgLysSerSerAlaGlnGlnAlaLeuHisArgSerAsnLys 365
 OY 1239 GATGTAATTTTGTGTCGCGAGAGCTCTCAAAATCCAGAGCAGAAACCATATGTTT 1298
 DB 366 AspGlySerPheLeuIleArgLysSerSerGlyHisAspSerLysGlnProTyrThrLeu 385
 OY 1299 GTGCTGTTTATGGCAACAGGTCTACAAATGGAATCCGTTTCCCTGAGAGAAATCAA 1358
 DB 386 ValValPhePheAsnLysArgValTyrAsnIleProValArgPheIleGluAlaThrLys 405
 OY 1359 CAGTTTGCCCTGGCGACAGACTACGAGAAATGAGATGTTGATCTGTGGAAGACATC 1418
 DB 406 GlnTyrAlaLeuGlyArgLysLysAsnGlyGlnGluTyrPheGlySerValAlaGluIle 425
 OY 1419 ATTGAACATACATATTTTCCCATTCCTGCTAAATAGATGGGAAGAC 1466
 DB 426 IleArgAsnHisGlnHisSerProLeuValLeuIleAspSerGlnAsn 441
 RESULT 9
 AAM59867
 ID AAM59867 standard; Protein; 457 AA.
 AC AAM59867;
 XX
 XX 20-NOV-1998 (first entry)
 XX
 XX Amino acid sequence of the mouse BLNK protein.
 KW Mouse; BLNK; B cells linker protein; apoptosis; Grb2; PLC-gamma;
 SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;

KW ras pathway; GDP; GTP; calcium pathway; antigen.
 XX Mus sp.
 XX W09832852-A1.
 XX
 XX 30-JUL-1998.
 XX
 XX 23-JAN-1998; 98MO-US01394.
 XX
 XX 17-MAR-1997; 97US-0819013.
 XX 24-JAN-1997; 97US-0788322.
 XX
 XX (UNIV) UNIV WASHINGTON.
 XX Chan AC, Fu C;
 XX
 XX WPI: 1998-427948/36.
 XX N-PSDB; AAV41903.
 XX
 XX Human B cell linker proteins - useful in the treatment of diseases
 XX involving increased or decreased apoptosis
 XX
 XX Disclosure: Fig 5; 56pp; English.
 XX
 CC This is the amino acid sequence of the murine BLNK (B cells linker
 CC protein) protein, used the treatment of diseases involving the
 CC increase and decrease of apoptosis, in the method of the invention.
 CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
 CC intracellular calcium levels and Grb2 is an adapter molecule
 CC containing two SH3 domains that mediate its interaction with the
 CC guanine nucleotide exchange factor, Son of Sevenless (Sos) which in
 CC turn activates the ras pathway by facilitating the exchange of GDP for
 CC GTP on the ras molecule. Activation of both ras and calcium pathways
 CC are required for efficient B cell antigen receptor function. BLNK
 CC binding proteins can be used to identify BLNK proteins in a target
 CC sample.
 CC
 XX
 XX
 SO Sequence 457 AA:
 Alignment Scores:
 Pred. No.: 2,036-17 Length: 457
 Score: 287.50 Matches: 106
 Percent Similarity: 38.31% Conservative: 53
 Best Local Similarity: 25.54% Mismatches: 139
 Query Match: 9,408 Indels: 117
 DB: 19 Gaps: 16
 US-09-856-061-1 (1-1721) x AAW59867 (1-457)
 QY 450 AACACGAGCTAGAGATCCGATCCAGCTGCGAAGCCATGCCATCAATGAATAAT 509
 DB 69 ASPTAATPTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 85
 QY 510 TTACGAGCAGACCTATCCAGATGCGAATATACAGATACACCTATTCCAGATATG 569
 DB 86 MetProAlaGluGluThrGlyAspSerSerTyrGlu 97
 QY 570 ATGAGAGCTCCCTCTGTGTACCTCCAGAGCTTGTGTCTCAGTGAAGACAAACGAG 629
 DB 98 -----ProProProAla-----GluGluGlnThrArg 106
 QY 630 GATGTGAGG-----ATGACACAGCTGGAAGAGAGTGGACAAACCTACCTCAAG 677
 DB 107 ValValIHisProAlaLeuProPheThrArgGlyGlyValAspAsnArgSer----- 124
 QY 678 GATGTGACAGCCACGCTTTAA-----GGATTCAAAATACAAAAAATAAACAGACT--- 731
 DB 125 -----SerGlnArgHisSerProProPheSerIleThrLeuProSerIleProSer 141
 QY 732 ---CCTTTGCCACCTCCCTGGCTGTATCATCTCTCCCT-----AAGAG 773
 DB 142 TrpProSerAlaLysAlaArgLeuAlaSerThrLeuProAlaProAsnSerLeuGlnLys 161

QY 774 TACCAACCTTACCCCGACACACCA-----GAGAGAGCATGCTACTTCGCT 824
 DB 162 ProGln---ValProLysProLysAspLeuLeuGlnAspGlnLysAspTyrValVal 180
 QY 825 CCAAGCCACCTTCCAGAAAGTCCAGAGGGGCCAGCAGACAGCT----- 872
 DB 181 ProValGluAspAsnAspIleAsnTyrIleHisProArgGluSerSerProProAla 200
 QY 873 -----GCAAAAGACTTCAAGAGGCTTGGAGCAGAAAGAAATCTCACCCAGACA 926
 DB 201 GluLysAlaProMetValAsnArgSerThrLysProAsnSerSerIleHisIleSer 220
 QY 927 AAGCCAGAACTCT----- 938
 DB 221 ProProGlyThrValAlaGlyArgAsnSerGlyValTrpAspSerLysSerIlePro 240
 QY 939 TCTTGCCCATCATCAACCAAAACACAGAGAGTCCA----- 977
 DB 241 AlaAlaProSerProLeuProAlaGlyArgAlaGlyLysLysProAlaThrProLeuLysThr 260
 QY 978 -----CTGCCATTGCCAGCTCTTCTACATGCCAGAAAGACAGATACACCCAGA 1031
 DB 261 ProValProLeuProLysProAsnAlaSerAsnValLysGluGluLysProValProAlaGlu 280
 QY 1032 GACCATACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
 DB 281 ArgHisArgGlySer-----SerHis 287
 QY 1092 -----AGCCCTGAATGCTGCCCTTGAATAACACAACTGGAGAA 1133
 DB 288 ArgGlnAspThrValGlnSerProValAlaPheProProGlnGlnLysProValHisGlnLys 307
 QY 1134 CTGACCCACACA----- 1145
 DB 308 ProValProLeuProAlaArgPheProGluAlaGlySerProAlaAlaAspGlyProPheHis 327
 QY 1146 -----AAGCCTGATGGAAGAGATGCTGCCAGATGAATG 1181
 DB 328 SerPheProPheAsnLeuThrPheAlaAspGlnGluGluLeuGlyLysProTrp 347
 QY 1182 TACATTTGAGATATACAGTCCGACGACGATGATGATGATGATGATGATGATGAT 1241
 DB 348 TyrAlaGlyAlaLysAspArgLysPheAlaGlnGluAlaIleHisArgSerAsnLysAsp 367
 QY 1242 GGTACTTTTGGTCCGAGCTGCTCTCAAAATCCAAAGCAGAACCTATGTTGGTG 1301
 DB 368 GlySerPheLeuIleArgLysSerPheGlyHisAspSerLysGlnProTyrThrLeuVal 387
 QY 1302 GTGTTTATGGGAACAGCTTACATGTGAATAATCCGTTCTCGAGAGCAATCAACAG 1361
 DB 388 AlaPhePheAsnLysArgValTyrAsnIleProValAlaPheIleGluAlaThrLysGln 407
 QY 1362 TTTGCCCTGGCAGACAGCTAGAGAAATGAGATGTTGATTTGTGGAGACATCAT 1421
 DB 408 TyrAlaLeuGlyLysLysAsnGlyGluGlyLysArgSerValAlaGluIleVal 427
 QY 1422 GAACACTACACATATTTCCATCTGCTATATAGATGGAAAGAC 1466
 DB 428 AsnSerHisGlnHisAsnProLeuValLeuIleAspSerGlnAsn 442
 RESULT 10
 AAW70587
 ID AAW70587 standard; Protein: 95 AA.
 AC AAW70587;
 XX
 XX 28-JAN-1999 (first entry)
 DE Human SLP-76 SH2 domain.
 KW Human: SLP-130; SLP-76; SH2 domain; leukocyte protein; allergy; cancer;
 KW phosphoprotein; autoimmune disease; T-cell antigen receptor; TCR;

DB 21 AsnphSerLeupProLysAsnArgSerTyrProArgIleAsnSerAlaThrGlyInTrp 40
 OY 375 CGAGGGGTCTGGAGCCACTTCGGATCAGCAAGAACTTGGCTGGGCTCCAGGTGA 434
 DB 41 GlnAlaGlyMetAsnLysProLeuLysPrtPgluArgAsnPhleAlaValLeuAspGly 60
 OY 435 GAAAGATGCAACAGTACACACGACTACGAAATCCTGAGTTCAGCTGCTGAAGCATGG 494
 DB 61 AlAlaGlyLysHLeSerAsp****AspAspProGluLeuArgMetGluGluThrTrp 80
 OY 495 CCATCAATGAAA 506
 DB 81 GlnSerIleLys 84
 RESULT 12
 ABB38147
 ID ABB38147 standard; Peptide: 46 AA.
 AC ABB38147;
 XX
 DT 04-FEB-2002 (first entry)
 DE Peptide #5653 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 PS Claim 27; SEQ ID NO 30782; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 46 AA;
 Alignment Scores:
 Pred. No.: 6.14e-11 Length: 46
 Score: 212.00 Matches: 41
 Percent Similarity: 91.30% Conservative: 1
 Best Local Similarity: 89.13% Mismatches: 4
 Query Match: 6.93% Indels: 0
 DB: 22 Gaps: 0

US-09-856-061-1 (1-1721) x ABB38147 (1-46)
 OY 1239 GATGGTACTTTTGGTCCGAGACTGCTACAAATCCAGCAACCATATGTTTG 1298
 DB 1 AspGlySerPhleLeuValArgAspCysSerThrLysSerLysGluGluProTyrValLeu 20
 OY 1299 GTGGTGTTTTGGGACACAGCTCTACATATGAAATCCGTTCTCGAGAGCATCA 1358
 DB 21 AlaValPheTyrGluAsnLysValTyrAsnValLysIleArgPhleLeuGluArgAsnGln 40
 OY 1359 CAGTTTGCCCTGGGCACA 1376
 DB 41 GlnPheAlaLeuGlyThr 46
 RESULT 13
 ABB23354
 ID ABB23354 standard; Protein: 46 AA.
 AC ABB23354;
 XX
 DT 23-JAN-2002 (first entry)
 DE Protein #5353 encoded by probe for measuring heart cell gene expression.
 XX
 KW Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488899/53.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX
 PS Claim 15; SEQ ID NO 25124; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21533-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarray.
 CC By measuring gene expression, the probes are useful for predicting
 CC diagnosing, grading, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 46 AA;
 Alignment Scores:
 Pred. No.: 6.14e-11 Length: 46
 Score: 212.00 Matches: 41

Query Match: 6.93% Indels: 0
DB: 22 Gaps: 0
US-09-856-061-1 (1-1721) x NAM71296 (1-46)
QY 1239 GATGGTACTTTTGTGTCGAGACTGCTCTACAAAATCCAGGACAGAACCATATGTTTG 1298
|||||:|||||
DB 1 AspGlySerPheLeuValArgAspCysSerThrLySerLySGluGluProTyrValLeu 20
QY 1299 GTGGTGTATTATGGGAACAAGGCTACATGTGAAAATCCGTTTCTCGAGACCAATCAA 1358
|||||
DB 21 AlaValPheTyrGluAsnLysValTyrAsnValLysIleArgPheLeuGluArgAsnGln 40
QY 1359 CAGTTTGCCCTGGGCACA 1376
|||||
DB 41 GlnPheAlaLeuGlyThr 46

Search completed: April 21, 2003, 12:20:45
Job time : 90.294 secs


```

; LENGTH: 456 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-819-013-1

```

Pred. No.:	2, 21e-22	Length:	466
Score:	319.00	Matches:	99
Percent Similarity:	39.658	Conservative:	58
Best Local Similarity:	25.008	Mismatches:	139
Query Match:	10.448	Indels:	100
DB:	2	Gaps:	12

QY	444	AAACGTTAACACAGCACTCTCGAAGAACTCCGAGTCCGCTGAAGGATGCCCATCAATG	503
Db	91	AsnAlaSprspsrSerTyrGIuPrProProValGIu-----GInGIuThrArg	106
QY	504	AAATTTTACCAGCCAGCACTTATCCAGAAATCGAAATACGACATACAGCTATTTCAG	563
Db	107	ProValHisProAlaLeuProPheAlaArgGIuGIuThrLeuAspAsnArgSerGIn	126
QY	564	GATATGATGAGAGCTCCC-----CTTCTGTACCTCCCAAGGCTTGTCTCCACT	614
Db	127	-----ArgHisSerProProPheSerLysThsLeuProSerLysProSerTTrProSer	144
QY	615	GAGAGCAAAACACAGGATGTGAGATGACACAGACTGGAAGAGTGCACAGCTTACTTC	674
Db	145	GIuLysLsAlaArgLeuThrSerThrLeuProAlaLeuThAlaLeuGIuLysProGIuVal	164
QY	675	AAGGATGTAGAGCCACACGCTTTAAAGATTCAATACACAAAATTAACAAAGACTCCT	734
Db	164	-----	164
QY	735	TTCGCACTTCCTGGCGCTGCTATCACTTCGCCAGAGATACCAACCTTACCCCGACA	794
Db	165	-----ProIrisProLysGIuLeuLeuGIuAspGIuAlaAspIryValValProVal	182
QY	795	CCACGAGGAGAGAGCTGCATATCTTCGCTCCAAAGCCCACTTTCCAGAGTCCAGAG	854
Db	183	GIuAspAsnAspGIuAsnTyrIleHisProThrGIuSerSerSerProProGIuLys	202
QY	855	GGGCCCCAGCAGAGAGTGCAAAAGACTTCAGAGGCTCTTGAGGACAAAGAAATCT	914
Db	203	AlaProMet-----Val	206
QY	915	CACCAACACAAAGCCAGAAATCTTCTGCCATCATCAACCAAAACACA-----	965
Db	207	AsnArgSerThrLysProAsnSerSerThrProAlaSerProProGIuThrLaseArgI	226
QY	966	-----CAGAAGAGTCCACCTGSCCATGGCCAGCTCTTCTTCAATG	1004
Db	227	ArgAsnSerGIuAlaTrpGIuThrLysSerProProPro--AlaAlaProSerProLeu	245
QY	1005	CCA-----GGAAGACAC-----AGTATCAACAGCCAGAGAC	1034
Db	246	ProArgAlaGIuLysLysProThrThrProLeuLysThrThrProValAlaSerGIuGIn	265
QY	1035	CATACAGCTAGCATG-----CAGCACTGTCTCTGCTCAGAGATGCCAAGCTTCA--	1082
Db	266	AsnAlaSerSerValCysGIuLysGIuLysProIleProIleArgIuGInHisArgLysSer	285
QY	1083	-----GCCAGCCACAGCCCTCGAATGCTGCCCTTGAAACAAACAACACTCGGAG	1130
Db	286	HisArgGIuGIuAlaValGIuLysSerProValPheProProAlaGIuLysGIuIleHisGIu	305
QY	1131	AAACCTGACCC-----	1142
Db	306	LysProIleProLeuProArgPheThrGIuGIuLysAsnProThrValaSpGIuProLeu	325

```

QY 1143 -----ACAAAGCGTCGATGACAAAGATGTCGTGGCAGAAATGA 1178
Db 326 ProliehesetserAsnserThrlIleSerGIuGlnIuNlaIagIlyallleucyslyspro 345
QY 1179 TGGTACATTGCGAAGATACAGTGGCCAGGACGTGAGAGATGCTTTAATGAAAGAACAAAG 1238
Db 346 TPTPTTlAagIlaCyAsaPaRgLyseRalagIuGlnIaleuNlaHsIserGAsnLys 365
QY 1239 GATGGACATTTTTGGTCCGAGACTGCTGCATACAAATCCAGAGCAACCATATGTTTTG 1298
Db 366 AAGGlyseRphenleuIlealGlySerSeRGIyHIsAspserLyseInProtyrThrlleu 385
QY 1299 ATGGGTGTTTATGGAGAACGTTACCAATGTGAAAAATCCGTTTCCTCGAGCAATCA 1358
Db 386 ValIalPhepheasnlYsArgValTyAsnIleProvalArgPheIleGlnIaleHrlyS 405
QY 1359 CAGTTTGGCCCGGCGACAGACTACGAGGAATGACATGTTGATTCTGTGGACATCATC 1418
Db 406 GlnTyAlaIlealGlyAlaGlyLysAsnGlyGlnIuTyPheGlySerTyAlaIleGlnIle 425
QY 1419 ATGACACACTACACATATTTTCCCATCTGCTAATGATGATGGAGAAAGAC 1466
Db 426 IlealGAsnHIsGlnHIsseRProleuValleuIleAspserGIaasn 441

RESULT 2
US-08-729-416C-1
/ Sequence 1, Application US/08729416C
/ Patent No. 6013767
/ GENERAL INFORMATION:
/ APPLICANT: NAKAMURA, TAKESHI
/ TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
/ TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
/ STREET: 1100 NEW YORK AVENUE, N.W.
/ CITY: WASHINGTON
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20005-3918
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/729,416C
/ FILING DATE: 11-OCT-1996
/ CLASSIFICATION: 336
/ ATTORNEY/AGENT INFORMATION:
/ NAME: PERRY, GLENN J.
/ REGISTRATION NUMBER: 28458
/ REFERENCE/DOCKET NUMBER: 7898/225948
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-861-3000
/ TELEFAX: 202-822-0944
/ TELEX: 6714627 CUSH
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 474 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-729-416C-1

Alignment Scores:
Pred. No.: 1,64e-07 Length: 474
Score: 165.50 Matches: 106
Percent Similarity: 36.40% Conservative: 60
Best Local Similarity: 23.25% Mismatches: 161
Query Match: 5.41% Indels: 129
DB: Gaps: 25

```

US-09-856-061-1 (1-1721) x US-08-729-416C-1 (1-474)

QY 291 GGATTCGGTGCATGTAGATTCAGAACGTCCTCTCTGTAATAATAGTCATGCCAAGC 350
 ||| :|||:||||| :|||:||||| :|||
 Db 94 GtlySerAspGlyAspPheLeuValArg-----LysSerThrIleAsnProGly 409
 ||| :|||:||||| :|||:||||| :|||
 QY 351 CTGAGCAGTCCCAAGGGCGCTGCGAGCGGCTTGGAAACCACTTCCGGATCCAGAAAG 410
 ||| :|||:||||| :|||:||||| :|||
 Db 114 LeuAsnLeuArgThrProAspSerLysGlnIleLeu-----AlaAsnHisMetArg 131
 ||| :|||:||||| :|||:||||| :|||
 QY 411 AACTGGCTGGGTCCAGCGGTGAGAAATAAGCAAGTAAACAAGTAC-----461
 ||| :|||:||||| :|||:||||| :|||
 Db 132 SerIleSerPheAlaSerGlyAspPro-----AspThrIleAspArgValAlaTyr 149
 ||| :|||:||||| :|||:||||| :|||
 QY 462 -----GAAATCTCT-----GAGTTCAGCTGCTGAGAGCATGCCATCAATG 503
 ||| :|||:||||| :|||:||||| :|||
 Db 150 ValAlaLysAspProValAsnArgArgAlaCysHisIleLeuGlnCysCysAspGly-- 168
 ||| :|||:||||| :|||:||||| :|||
 QY 504 AAATTTTACCAGCAGACCTATCCAGATCCGGAATACAGATACAGCTATTTCAG 563
 ||| :|||:||||| :|||:||||| :|||
 Db 169 -----LeuAlaGlnAspValIleGlySerIleGlyGlnAlaPheGlnLeuArgPheLys 166
 ||| :|||:||||| :|||:||||| :|||
 QY 564 GATATGATGAGGCTCCCTCTGTGTACTCCCAAGGCTTCTGTCTCCATGAGAGCAA 623
 ||| :|||:||||| :|||:||||| :|||
 Db 187 GltArgLeuGlnCysProThrLysIlePro-----AlaLeuHisAspArgMetGln 203
 ||| :|||:||||| :|||:||||| :|||
 QY 624 ACCGAGATGTGAGATGACACAGCTGAGAAAGTGGACAAAGCTTACCTTCAGATGTC 683
 ||| :|||:||||| :|||:||||| :|||
 Db 204 SerLeuAspGluProThrProGlu--GlnGlnIleLysP-----215
 ||| :|||:||||| :|||:||||| :|||
 QY 684 AGAAGCAAGCGCTTTAAAGATTCAAATACAAATAAACAAGATCCCTTGGCACT 743
 ||| :|||:||||| :|||:||||| :|||
 Db 216 -----GlySerAspHisProGlyTyrIleAsnSerIleProSerLysMet 229
 ||| :|||:||||| :|||:||||| :|||
 QY 744 CTTGGCTGCTATCACTCTCCCAAGAGTACCACCTTACCCCAAGACACCAAG 803
 ||| :|||:||||| :|||:||||| :|||
 Db 230 ProProGluGlyPheLeuAspThrArgLeuLysProArgProHisAlaPro-----247
 ||| :|||:||||| :|||:||||| :|||
 QY 804 GAGAGCAGTGCATCTTCGCTCCCAAG-----830
 ||| :|||:||||| :|||:||||| :|||
 Db 248 ---AspThrIleGlnIlePheAlaGlyLysGlnIleThrTyrTyrGlnIleArgHisLeuGly 266
 ||| :|||:||||| :|||:||||| :|||
 QY 831 CCCACCTTTCAGAA---GTCCAGAGGGGGGCGCAGCAGAGAGTGCAGAAAGACTTCAGT 887
 ||| :|||:||||| :|||:||||| :|||
 Db 267 AspThrPheGlyGlnAspTyrGlnIleThrProLeuArgGlnGlySerAsp-----284
 ||| :|||:||||| :|||:||||| :|||
 QY 948 TCATCAAAACCAAAACACACAGAGTCCAGCTGCCATTCGCACTCTCTCATATGCA 1007
 ||| :|||:||||| :|||:||||| :|||
 Db 303 ---ThrTyrValAsnThrGlnIleLeuPro-----312
 ||| :|||:||||| :|||:||||| :|||
 QY 1008 GGAAGACACAGTATACAGCCAGAGACCATACAGGTACATGACAGCTGCTGCTAG 1067
 ||| :|||:||||| :|||:||||| :|||
 Db 313 -----GlnAlaTyrProAla---317
 ||| :|||:||||| :|||:||||| :|||
 QY 1068 AGATGCCAAGCTGCAGCCAGCAGCCCTCGA-----ATGCTGCCCTAT 1112
 ||| :|||:||||| :|||:||||| :|||
 Db 318 ---AlaValSerSerAlaGlnIleSerSerProArgLysAspLeuPheAspMetLysProPhe 336
 ||| :|||:||||| :|||:||||| :|||
 QY 1113 GAAACACAAACTCGGAGAAACT-----1136
 ||| :|||:||||| :|||:||||| :|||
 Db 337 GlnAspAlaLeuLysAsnGlnProLeuGlyProValLeuSerLysAlaAlaSerValGlu 356
 ||| :|||:||||| :|||:||||| :|||
 QY 1137 -----GACCCCAACAAAGCTGATGAGAAG-----GATGCTGCGAG 1172
 ||| :|||:||||| :|||:||||| :|||
 Db 357 CysIleSerProValSerProArgLysProAspAlaLysMetLeuGlnGluLeuGlnAla 376
 ||| :|||:||||| :|||:||||| :|||
 QY 1173 AATGAATGGTACATTGAGATACAGTCCGACAGTGGCAAGTGTGTAATGAAGAG 1232
 ||| :|||:||||| :|||:||||| :|||

Db 377 GltThrTyrGlnGlyGlnMetSerArgLysGlnAlaGlnGlyLeuLeu-----393
 ||| :|||:||||| :|||:||||| :|||
 QY 1233 AACAGAGATGATCTTTTGGTCCAGACTCTCTACAAATCCAGACAGACCA---1289
 ||| :|||:||||| :|||:||||| :|||
 Db 394 GtlyLysAspGlyAspPheLeuValArg-----LysSerThrIleAsnProGly 409
 ||| :|||:||||| :|||:||||| :|||
 QY 1290 ---TATGTTTGGTGGTCTTTATGGCAAGGTCCTACATGTGAAATCCGTTTCTC 1346
 ||| :|||:||||| :|||:||||| :|||
 Db 410 SerPheValLeuThrGlyMetHisAsnGlnAlaIleHisIleLeuLeuValAspPro 429
 ||| :|||:||||| :|||:||||| :|||
 QY 1347 GAGAGCAATCAACAGTTTCCCTGGGCAAGACTACGAGGAATGAGATGTTTATCT 1406
 ||| :|||:||||| :|||:||||| :|||
 Db 430 Gln-----GlyThrIleArgThrLysAspArgValPheAspSer 442
 ||| :|||:||||| :|||:||||| :|||
 QY 1407 GTGAGACATCATTCATGACACTAC-----ACATATTTCCATTCCTGTAATAGATGG 1460
 ||| :|||:||||| :|||:||||| :|||
 Db 443 IleSerHisLeuIleAsnHisHisLeuGlnIleSerLeuProIleVal-----458
 ||| :|||:||||| :|||:||||| :|||
 QY 1461 AAGACAGCTGACGACGAGAAACAGTCACTACCCAGCCACTG 1508
 ||| :|||:||||| :|||:||||| :|||
 Db 459 -----SerAlaGlySerGlnLeuGlnIleProVal 470
 ||| :|||:||||| :|||:||||| :|||

RESULT 3

US-08-729-416C-7
 ; Sequence 7, Application US/08729416C
 ; Patent No. 6013767
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAMURA, TAKESHI
 ; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
 ; TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
 ; STREET: 1100 NEW YORK AVENUE, N.W.
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3918
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/729,416C
 ; FILING DATE: 11-OCT-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PERRY, GLENN J.
 ; REGISTRATION NUMBER: 28458
 ; REFERENCE/DOCKET NUMBER: 7898/225948
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-861-3000
 ; TELEFAX: 202-822-0944
 ; TELEX: 6714627 CUSH
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 594 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-729-416C-7

Alignment Scores:

Pred. No.:	1,84e-07	Length:	594
Score:	165.50	Matches:	106
Percent Similarity:	36.40%	Conservative:	60
Best Local Similarity:	23.25%	Mismatches:	161
Query Match:	5.41%	Indels:	129
DB:	3	Gaps:	25

US-09-856-061-1 (1-1721) x US-08-729-416C-7 (1-594)

```

QY 291 GCATTCGGTATGATTCAGAACTCTCTGCTGAATAATAGTCATGCCAAC 350
   ||| :|||:||||| :|||:||||| :|||
Db 214 GYLysSerAsnLeuGlnPheAlaGlyMetSerIleSerLeuThrIleSerThrAlaSer 233
QY 351 CTGACGACGTGCCAAGAGGGGGTGTGACGGGTTCTGAAACCTTCGGATCAGCAAG 410
   |||:||||| :|||:||||| :|||:||||| :|||
Db 234 LeuAspLeuArgThrProAspSerLysGlnIleIle-----AlaAsnHisMetAla 251
QY 411 AACTGGCTGGGCTCCCGATGGAGAAATATGCAACAGTACACACACTAC----- 461
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 252 SerIleSerPheAlaSerGlyLysPro-----AspThrThrAspTyrValAlaTyr 269
QY 462 -----GAAGATCCT-----GAGTTCACAGCTCTGAAAGGCATGCCATCAATG 503
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 270 ValAlaLysAspProValAsnArgArgAlaCysHisIleLeuGlnCysCysAspGly---- 288
QY 504 AAAATTTTACACCCGACACTATCCGAAATGGAATACGCAAGATACACGATATTTCAG 563
   ||| :||| :||| :||| :||| :|||
Db 289 -----LeuAlaGlnAspValIleGlySerIleGlyGlnAlaPheGlnLeuArgPheLys 306
QY 564 GATATGATGAGGCTCCCTCTGTACTCCCAAGGCTTGTCTCCATGACAGACAA 623
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 307 GlnTyrLeuGlnCysProThrIleSerIlePro-----AlaLeuHisAspArgMetGln 323
QY 624 ACCAGGATGTGAGATGACACAGCTGGAAGAAGTGGACAGACCTTCAAGAGATGC 683
   :||| :||| :||| :||| :||| :|||
Db 324 SerLeuAspGlnProThrPheGlu---GluGlnGlyAsp----- 335
QY 684 ACAAGCCACGCTTAAAGATTTCAAAATACACAAAATTAACAAGACCTCTTGCACCT 743
   ||| :||| :||| :||| :||| :|||
Db 336 -----GlySerAspHisProTyrTyrAsnSerIleProSerLysMet 349
QY 744 CCTCGGCTGCTATACACTCTCCCAAGAAATACACACCTTACCCCAACACACAGAG 803
   ||| :||| :||| :||| :||| :|||
Db 350 ProProProGlnGlyPheLeuAspThrArgLeuLysProArgProHisAlaPro----- 367
QY 804 GAGACAGTGCCTACTGCTGCCAAG----- 830
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 368 ---AspThrAlaGlnPheAlaGlySerGlnThrTyrTyrGlnGlyArgHisLeuGly 386
QY 831 CCCACCTTTCAGAA---GTCCAGAGGGGGCCCGACAGAGAGTGCACAAAGCTTCACT 887
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 387 AspThrPheGlnGlyAspTyrPheGlnThrProLeuArgGlnGlySerSerAsp----- 404
QY 888 AGGCTCTTGGAGCAGAAAGATCTACACACACAGACAAAGCCGAATCTTCTGCCCA 947
   :||| :||| :||| :||| :||| :|||
Db 405 -----IleTyrSerThrProGlnGlyLysLeuHisValAlaProThrGlnGlyAlaPro 422
QY 948 TCATCAACCAACAAACACAGAGAGTCCACCTGCCATTCGCAGCTCTTCTACATGCCA 1007
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 423 ---ThrTyrValAsnThrGlnGlnIleProPro----- 432
QY 1008 GGAAGACACTATACAGCCAGACCATACAGGTACATGACAGACACTGTCCTGCTAG 1067
   ||| :||| :||| :||| :||| :|||
Db 433 -----GlnAlaTyrProAla----- 437
QY 1068 AGATGCCAAGCTGACGCCAGCCACAGCCCTCGA-----ATGCTGCCCTAT 1112
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 438 ---AlaValSerSerAlaGlnSerSerProArgLysAspLeuPheAspMetLysProPhe 456
QY 1113 GAAAACCAAACTCGAGAAACT----- 1136
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 457 GluAspAlaLeuLysAsnGlnProLeuGlnYrProValLeuSerLysAlaAlaSerValGln 476
QY 1137 -----GACCCCAACAAAGCTGATGAGAG-----GATGTCTGGCAG 1172
   :||| :||| :||| :||| :||| :|||
Db 477 CysIleSerProValSerProArgAlaProAspAlaLysMetLeuGlnGlnAla 496
QY 1173 AATGAAGGTGCTGATGAGATTAAGTCAGCCAGCAGCTGGAAGTGTGTTAATGAAGAG 1232
   ||||| :||| :||| :||| :||| :|||
Db 497 GluThrTyrPyrGlnGlyGlnMetSerArgLysGlnAlaGlnGlyLeuLeu----- 513
QY 1233 AACAAGATGCTACTTTTGGTCCGAGACCTGCTCAAAAATCCAAAGGCAAGAACCA--- 1289

```

```

Db 514 GYLysAspGlyAspPheLeuValArg-----LysSerThrThrAsnProGly 529
QY 1290 ---TATGTTTGTGCTGTGTTTATGGAACAGGCTCTACAAATGTAATAATCCGTTCTC 1346
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 530 SerPheValLeuThrGlyMetHisAsnGlnAlaLysHisLeuLeuValAspPro 549
QY 1347 GAGACCAATCAACACTTTCCTGGGACAGGACTACGAGAAATGAGATGTTGATTC 1406
   ||| :||| :||| :||| :||| :|||
Db 550 Glu-----GlyThrIleArgThrLysAspArgValPheAspSer 562
QY 1407 GCGAAGACATCATGAACTAC-----ACATATTTTCCCATCTCTCTAATGATCGG 1460
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 563 IleSerHisLeuIleAsnHisHisLeuGlnSerLeuProIleVal----- 578
QY 1461 AAAGCAAGCTGCACGCGAGAAACAGTCTACTACCCACCGACACTG 1508
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 579 -----SerAlaGlySerGlnLeuLys---LeuGlnGlnProVal 590

```

```

RESULT 4
US-08-807-342B-4
; Sequence 4, Application US/08807342B
; Patent No. 6077686
; GENERAL INFORMATION:
; APPLICANT: Der. Channing
; APPLICANT: O'Bryan, John P.
; TITLE OF INVENTION: No. 6077686el SHC Proteins
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: Howson and Howson
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,342B
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,516
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: MTS20USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-807-342B-4

```

```

Alignment Scores:
Pred. No.: 3 33e-05 Length: 550
Score: 142.00 Matches: 112
Percent Similarity: 34.16% Conservative: 67
Best Local Similarity: 21.37% Mismatches: 203
Query Match: 4.65% Indels: 142
DB: 3 Gaps: 23

```

US-09-856-061-1 (1-1721) x US-08-807-342B-4 (1-550)

1053 CACTGTCTGCTCAGAGATGCCAAGCTGCAGCCACAGCCCTCGAATGCTGCCCTAT 1112

Alignment Scores:

```

Pred. No.: 4,286-05 Length: 469
Score: 140.50 Matches: 101
Percent Similarity: 34.59% Conserved: 55
Best Local Similarity: 22.39% Mismatches: 196
Query Match: 4,608 Indels: 99
DB: 1 Gaps: 19

US-09-856-061-1 (1-1721) x US-08-353-550-6 (1-469)

OY 243 GTCCAGAGCNCATGACACCCAGGCAATTAAGACAAAGAAAGAA----- 290
DB 77 ValProGlyValArgGlySerTrpLysLysAlaProAsnLysAlaLeuAlaSerVal 96
OY 291 ---GGATTCGGTATGTGAGATTCAGAAACGCTCTCTGCTGTAATAATAGTCATGCCA 347
DB 97 LeuGlyLysSerAsnLeuArgPheAlaGlyMetSerLysSerLysLysLeuThrAsp 116
OY 348 AGCCTGAGCGTGCAGAAAGGCGGTGCGAGGCTTGTGAACTCCGATCAGAGA 407
DB 117 GlyLeuSerLeuSerValProAlaThrArgGlnValIle-----AlaAsnHisHisMet 134
OY 408 AGGAATTGGCTGGGTCCAGGTGGAGAAATGCAACAGTAACAAGACTAC----- 461
DB 135 ProSerLysSerPheAlaSerGlyGlyAsp-----ThrAspMetThrAspTyrValAla 152
OY 462 -----GAAGTCCTGAGTCCAG-----CTGCTGAAGGCGATGCCATCA 500
DB 153 TyrValAlaLysAspProIleAsnGlnArgAlaCysHisLysLeuLeuGlyCysGlyGly 172
OY 501 ATGAAATTTTACACGACGACCATTCAGGAATCG-----GAATTCGAGATATACAGC 554
DB 173 Leu-----AlaGlnSerLysLysLeuSerThrValGlyGlnAlaPheGlyLeuArg 188
OY 555 TATTCAGAGATATGATGAGAGGCTCC---CTTCGTACCTCCCAAGCTTCTGTCTCC 611
DB 189 PheLysGlnTyrLeuHisSerProProLysValAlaLeuProGluArgLeuAlaGly 208
OY 612 ACTGAGAGCAAAACAGGAGTGTAGATGACACAGCTGAA----- 653
DB 209 ProGluLysSerAlaTrpGlyAspGlyLysPheLysSerLeuGlnHisAsnTyrTyrAsnSer 228
OY 654 -----GAAGTGCACAAAGCTTACCTTCAGAGATGTCAGAACGCCAAGCTTTAAA 701
DB 229 IleProGlyLysGluProProLeuGlyGlyLeuValAspSerArgLeuAlaLeuThrGln 248
OY 702 GGATTCAAATACACAAATTAACAGACTCTTGGCCACCTCGGCGCTGATACACT 761
DB 249 ProCysAlaLeuThrAlaLeuAspGlnGlyProSerProSerLeuArgAspAlaCysSer 268
OY 762 CTCCCAAGAGTAGCAACACCTTACCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 821
DB 269 LeuProTrpAspValGlySerThrGlyThrAlaProProGlyAspGly----- 284
OY 822 GCTCCAAAGCCCACTTTCAGAAAGTCCAG-----AGGGGCGCCAGGCGAGAGAGT 872
DB 285 -----TyrValGlnAlaAspAlaArgGlyPro----- 293
OY 873 GCAAAAGACTCAGTAGGGCTCTTGAGACAGAAAGAAATGTCACACACAGCAAAAGCA 932
DB 294 ---ProAspHisGlnGlnHisLeuTyrValAsnThrGlnGlyLeuAspAlaProGluPro 312
OY 933 GAATTCCTTGGCCCATCATCAAAACCAAAACACAGAGAGTAGTCACATGCCATGCCAGC 992
DB 313 GluAspSer----- 315
OY 993 TCTTCTACATGCCAGAAAGCAGATATACAGCCAGAGACATACAGGTAGCATGCAG 1052
DB 316 -----ProLysLysAspLeuPheAspMetArgProPheGluAspAlaLeuAsp 331
OY 1053 CACTCTCTGCTCAGAGATGCCAGCTGACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1112
DB 332 -----LeuHisGlyLysSerValAlaAlaGlyValThrAlaAlaProLeuProLeu 348

```

```

OY 1113 GAAACACAAACTCGAGAAACCTGAC-----CCCAAAAGCTGATGAGAGAGAT 1163
DB 349 GluAspGlnTrpProSerProProThrArgArgAlaProValAlaProThrGluGln 368
OY 1164 GTCTGCAGAAATGAATGTGATCATTGAGAAATACAGTCCGACGAGCAGTGAAGATGTGTA 1223
DB 369 LeuArgGlnGluProTrpTyrHisGlyArgMetSerArgAlaAlaGlyArgMetLeu 388
OY 1224 ATGAAAGAGAACAGAGATGCTTTTGTGTCGAGACAGTGTCTGTAATAATCCAAAGCA 1283
DB 389 -----ArgAlaAspLysPheLeuValArgAspSerValThr-----AsnPro 403
OY 1284 GAACATATCTTTGTGCTGTTTATGCGAACAGGCTACATGTGAAATCCGTTTC 1343
DB 404 GlyGlnTyrValLeuThrGlyMetHisAlaGlyGlnProLysHisLeuLeuValAsp 423
OY 1344 CTCGAGAGCAATCAACAGTTGCCCTGGGCACAGACTACGAGAAATGATGTTGAT 1403
DB 424 ProGlu-----GlyValValArgThrLysAspValLeuPheGln 436
OY 1404 TCTGTGAGACATCATGTGACACTACATATTTCCCATTCGTGATATGATGGAGAA 1463
DB 437 SerLysSerHisLeuLysPheHis-----LeuGlnAsnGlyGln 450
OY 1464 GACAAAGCTGCACGAGAAACAGTCTACTCTC 1496
DB 451 ProIleValAlaAlaGlnSerGlyLeuHisLeu 461

```

```

RESULT 7
US-08-551-687-6
; Sequence 6, Application US/08551687
; Patent No. 5925547
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Cavanaugh, William M.
; TITLE OF INVENTION: No. 5925547el protein Domain Which Binds
; TITLE OF INVENTION: Tyrosine Phosphorylated Proteins
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,687
; FILING DATE: 01-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/353,550
; FILING DATE: 09-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Matthew B.
; REGISTRATION NUMBER: P39,787
; REFERENCE/DOCKET NUMBER: 2307K-5731
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SFO ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-551-687-6

```

Alignment Scores:

Pred. No.	4	28e-05	Length:	46
Score:	140.50		Matches:	10
Percent Similarity:	34.59%		Conservative:	5
Best Local Similarity:	22.39%		Mismatches:	15
Query Match:	4.60%		Indels:	99
DB:	2		Gaps:	19

QY	243	GTGCGCAAGGACCCCTGACCCAGAGGCGCATATAAGCAACAGAAAGAA-----	290
Db	77	ValProGlyValAlaArgLysSerTrpLysLysAlaProAlaLysAlaLeuAlaSerVal	96
QY	291	----GGATTGGGATGTGATGATTCCAGAACSTCTCTGCTGAAATAATGTCATGGCCA	347
Db	97	LeuGlyLysSerAlaLeuArgPheAlaGlyMetSerLleSerLleHisLleSerThrAsp	116
QY	348	AGCCATGACAGATGCGCAAAAGGCGGTGTGAGCGGGTTCGGAACCACTTCGCGATCACGA	407
Db	117	GlyLeuSerLeuSerValProAlaIleThrArgGlnValIle-----AlaAsnHisMet	134
QY	408	AGGAACTTGGCTGGGGTCCCAAGTGGAGAAAAATGCACAAGTAAACAGACTAC-----	461
Db	135	ProSerLleSerPheAlaSerGlyLysPsr-----ThrAspMetHisArgTyrValAla	152
QY	462	-----GAAGATCCCTGAGTTCCAG-----CTGCTGAAAGCGATGGCCATCA	500
Db	153	TyrValAlaAlaLysPsrProLleAsnIleThrAlaCysHisLleLeuGlyCysGlyGly	172
QY	501	ATGAAATTTTACAGCCAGACCACTTCCAGGAATCG-----GAATACGCGAGATACAGC	554
Db	173	Leu-----AlaGlnSerLleLleSerThrValGlyGlnAlaPheGlyLeuArg	188
QY	555	TATTTCCAGGATATGATGGAGGCTCC---CTTCTGTACTCTCCCAAGCTTGTCTCC	611
Db	189	PheLysGlnTyrLleHisSerTrpProLysValAlaLeuProProGlyAlaGlyLeuAlaGly	208
QY	612	ACTGAGAGACAACACAGGATGTGAGATGCACACACTGGAA-----	653
Db	209	ProGlnGlySerAlaIleArgLysPheLysPsrLeuGlnHisAsnTyrTyrAsnSer	228
QY	654	-----GAAGTGGACAAGCCATCTTCAAGGATGTCAAGAACCCAAAGCTTTAA	701
Db	229	IleProGlyLysGlyProToroLeuGlyLysValLysPsrArgLeuAlaLeuThrGln	248
QY	702	GGATTCAATCAACAAATAAACAAAGACTCTTGGCCACTCTGGCGGCTGGTACACT	761
Db	249	ProCysAlaLeuThrAlaLeuAspGlnGlyProSerTrpSerLeuArgAspAlaCysSer	268
QY	762	CTCCCAAGAAATGACCAACCTTACCCCAAGCACACACAGAGAGAGACAGTGCATCTTC	821
Db	269	LeuProTrpAspValGlySerThrGlyThrAlaProGlyLysPsrGly-----	284
QY	822	GCTCCAAAGCCACACTTCCGAAGTCCAG-----AGGGGGCCCAAGCAAGAGAT	872
Db	285	-----TyrValGlnAlaAspAlaArgGlyPro-----	293
QY	873	GCAAAAGACTTCACTAGCGGCTTGGAGCGAAGAAGATCTCCACACAGCAACAGCA	932
Db	294	---ProAspHisGlyGlnHisLleTyrValAlaThrGlnGlyLysAspAlaProGlyPro	312
QY	933	GAATCTTCTTCCATCATCATGAAACCAAAACACACAGAAAGTCCACTGCTATGCGAC	992
Db	313	GlnAspSer-----	315
QY	993	TCTTCTACATGCCAGGAAGCACAGTATCAAGCCAGAGACCATACAGTACGATGGAG	105
Db	316	-----ProLysLysAspLeuPheAspMetArgProPheGlnAspAlaLeuLys	331
QY	1053	CACGTCTCTGCTCAAGATGCCCAAGCTGCAAGCCACAGCCACAGCCCTCGAATCTGCGCTAT	1112
Db	332	-----LeuHisGlyCysSerValAlaIleAlaGlyValThrAlaAlaProLeuProLeu	348

DQ
YQ 1113 GAAACACCAACTCGGAGAAACCTGCAG-----CCACAAGAAGCTCATGGAGAAGAT 1165
 ||||| ||| ||| ||| |||
Db 349 GUSPGRGLNTPRPSerProPofHrAlaArgAlaProValAlaLPrOfHstLUcLn 368

YQ 1164 CTCGGCAGAAATGAATGTACATTGGAGAAATACAGTCGCCAGGACAGTGGAMGATGTGTA 1223
 ::: ||| ||||| ||| ||||| ||| |||
Db 369 LeuArgInglInUProTPRtYnHSGLyArgMetSerArgAlaLeuAlaGlnArgMetLeu 388

YQ 1224 ATGAAGAGAACAGATGTGTTCTTTTGTCGCCAGATGCTCTTACAAAATCCAAAGCA 1283
 ||| ||||| ||||| ||||| |||
Db 389 -----ArgAlaAspGlyAspPheLeuValAlaTrpAspSerValThr-----AsnPro 403

YQ 1284 GAACCATATGTTTTGGTGGCTGTTTATTTGGAACAGGTCTACAAATGTGMAAATCCGCTTC 1343
 ||||| ||||| ::: ::: ::: :::
Db 404 GlyIleTyValLeuThrIleGlyMetHisAlaGlyIleProLyHisLeuLeuValAsp 423

YQ 1344 CTCGAGACCAATTCACAGTGGCCCTGGCACAGCACTACGAGAAATGAGATGTTGAT 1403
 ||| ||| ||| ||| ::: :::
Db 424 ProGU-----GlyValValArgThrLyAspValLeuPheLU 436

YQ 1404 TCCTGGAAAGCATCATTTGAACACTACACATATTTTCCCATTCTGCTTAATGATGGAAA 1463
 ||||| :::||||| ||||| ||| :::|||||
Db 437 SerIleSerHisLeuIleAspHisLS-----LeuGlnAsnGlyLn 450

YQ 1464 GACAAAGCTGCACGACGAGAAACAGTCTACCTC 1496
 ||| ::: ::: |||
Db 451 ProIleValAlaAlaGluSerGluLeuHisIleU 461

RESULT 8
US-08-72

Sequence 11, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENES
TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/7729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-11

Alignment Scores:	8.41e-05	Length:	474
Pred. No.:	137.50	Matches:	95
Score:	37.67%	Conservative:	70
Percent Similarity:			


```

Db      214 GlySerSerLeuGlnPheAlaGlyMetSerLeuThrIleSerThrAlaSer 233
QY      351 CTGAGAGTGCACAAAGGGCGGTCTGAGGCGTCTGGAAACCACTCCGATCCAGAAAG 410
Db      234 LeuAsnLeuArgThrProAspSerLysGlnIleIleSer-----AsnHisMetArg 251
QY      411 AACCTGGCTGGGTCCAGGTGCAGAAATAAGCAACAGTAAACAGCACTAC-----461
Db      252 SerLeuSerPheAlaSerGlyGlyAspPro-----AspThrThrAspIleValAlaTyr 269
QY      462 -----GAGATCCCT-----GAGTTCACAGCTCTGGAAGCATGGCCATCAATG 503
Db      270 ValAlaLeuAspProValAsnArgArgAlaCysHisIleLeuGlnCysAspGly---288
QY      504 AAAATTTTACAGCCAGACCTATCCAGAAATGGAAATAGCAGATACAGCTATTTCGAG 563
Db      289 -----LeuAlaGlnAspValIleGlySerIleGlyGlnAlaPheGlnLeuArgPheLys 306
QY      564 GATAGATGAGAGGCTCCCTCTGTACTCCCAAGGCTTGTCTGTCTCCAGTACAGACAA 623
Db      307 GlnTyrLeuGlnCysProSerLysIlePro-----AlaLeuGlnAspArgMetGln 323
QY      624 ACCAGGATGTAGATGATGACAGCTGACAGTGCAGAAAGTGCAGAACCTTCAAGATGTC 683
Db      324 SerLeuAspGlnProTrpThrGlu---GlnGlnGlyAspGlyPro-----337
QY      684 AGAAGCCAAAGCTTTAAAGATTCAAATACACAAATAACAGACTCTTGGCAGCT 743
Db      338 -----AspHisProTyrTyrAsnSerValProAsnLysMet 349
QY      744 CCTCGGCTGTACTACTCTCCCAAGAAAGTACCAACCTTACCCCAAGCAGCAGAG 803
Db      350 PropProGlyGlyPheLeuAspAlaArgLeuAlaArgProHisAlaPro-----367
QY      804 GAGACAGTGCCTACTCTGCTCCAAAGGCCACCTT-----839
Db      368 AspAlaIleAlaGlnPheSerGlyLysGlnIleThrTyrTyrGlnGlyArgHisLeuGlyAsp 387
QY      840 -----CCAGAGTCCAGAGGGGCCAGGAGAGAGTGCAGAAAGTTCAGTAGG 890
Db      388 AlaPheGlyGlnAspTrpGlnArgAlaProThrArgGlnGlySerLeuAsp-----404
QY      891 GTCCTGGAGCAGAGAGAGATCTCACACACAGACAAAGCCAGAAATCTTCTGGCCATCA 950
Db      405 IleTyrSerThrProGlnGlyLysAlaHisMetVal---ProValGlyGlnThrPro---422
QY      951 TCAAACCAAAACACACAGAGAGTCCACCT-----GCCATTGCCAGCTCT 995
Db      423 ThrTyrValAsnThrGlnProValProGlnValTrpProAlaIleAlaThrSerThr 442
QY      996 TCCATACATGCCAGAAAGACACAGATACACACAGACCATACAGTAGTACGATGCAGAC 1055
Db      443 GluSerSerProArgLysAspLeuPheAspMetLysProPheGlnAspAlaLeuArgAsn 462
QY      1056 TGTCTGCTCAGAGATGCAGAGTGCAGAGCCACAGACCCCTGCAATGCTCCCTATGAA 1115
Db      463 GlnProLeuGlyProValLeuSerLysAlaIleSerValGlnCysIleSerProVal---481
QY      1116 AACCAAACTCGAGAAACTGACCCCAACAAAGCCTGATGAGAG-----GATGTC 1166
Db      482 -----ThrProArgAlaProAspAlaLysMetLeuGlnGlnLeu 494
QY      1167 TGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1226
Db      495 AsnAlaGlnProTyrTyrGlnGlyGlnMetSerArgLysGlnAlaGlnAlaLeuLeu---513
QY      1227 AAAGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1286
Db      514 -----GlnGlnAspGlyAspPheLeuValArg-----LysSerThrThrAsn 527
QY      1287 CCA-----TATGTTTGGTGGTGTGTTTATGGAGCAAGCTTACAAATGCAAAATCGGT 1340

```

```

Db      528 ProGlySerPheValLeuThrClyMetHisAsnGlyGlnAlaIleLysHisLeuLeuVal 547
QY      1341 TTCCTGAGAGCAATCAACAGTTGGCCCTGGCCAGAGCACTAGAGAAATGATGTTT 1400
Db      548 AspProGlu-----GlyThrValArgThrLysAspArgValPhe 560
QY      1401 GATTCGTGTGGAAGACATCATTTACACACTAC-----ACATATTTTCCATCTGCTAATA 1454
Db      561 AspSerIleSerHisLeuIleThrTyrHisLeuGlnSerSerLeuProIleVal-----578
QY      1455 GATGGAAAGACAGAGTCTCAGCAGAGAAACAGTCACTACCCAGCAGCACTG 1508
Db      579 -----SerIleLysSerGlnLeuGlyCys---LeuArgGlnProVal 590

RESULT 10
US-08-434-730-16
; Sequence 16, Application US/08434730
; Patent No. 5637463
; GENERAL INFORMATION:
; APPLICANT: Dalton, Stephen
; APPLICANT: Kochan, Jarema P
; APPLICANT: Osborne, Mark A
; TITLE OF INVENTION: METHOD TO DETECT PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTIONS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: USA
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,730
; FILING DATE: 04-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Semionow, Raina
; REGISTRATION NUMBER: 39022
; REFERENCE/DOCKET NUMBER: 9069
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)235-4391
; TELEFAX: (201)235-2363
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 756 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-434-730-16

Alignment Scores:
Pred. No.: 0 00021
Score: 134.50 Length: 756
Percent Similarity: 33.178 Matches: 128
Best Local Similarity: 21.444 Conservative: 70
Query Match: 4.408 Mismatches: 204
DB: 1 Indels: 195
Gaps: 29

US-09-856-061-1 (1-1721) x US-08-434-730-16 (1-756)
QY      37 TCGAAACCAAACTCAACAGC-----ACATACAGCAGCAGCTCTGCTGAA 84
Db      197 AsnSerAsnSerSerSerGlyAlaGlyThrValGlyArgAlaLeuAlaAsnAsp 216
QY      85 GGCCTCTGCTGGAGGAGAGACATGATCAACTTATCTTACAGAGTGTCCAGATGCA 144

```


TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 670 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-09-856-061-2

Alignment Scores:

Pred. No.: 0.00022 Length: 670
 Score: 134.00 Matches: 120
 Percent Similarity: 33.71% Conservative: 60
 Best Local Similarity: 22.47% Mismatches: 180
 Query Match: 4.38% Indels: 174
 Gaps: 27

US-09-856-061-1 (1-1721) x US-08-980-080-2 (1-670)

```

QY 37 TCGAAGAACCAAACTACACAGC-----ACATACAGGCACTCTCTGTGMA 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 AanserAmsrAmsrAmsrserelgylalaglythrvallgyllysallaleuAlasnspp 216
QY 85 GGACTCTGCTGAGGGAGAGAACATGTCACTCTTACAGAGTGTCTCAGATGCGA 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 GlyThrSerProglYglunary----- 223
QY 145 CCGTGACCCCTTCACAGAGCTAGCCGTCTACACACTGAGCCCTTGACTAAAGAGA 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 ---Trp----- 232
QY 205 CTGACAGGCTGAGTTGAAGATCCCTCTTTGCCAGGTGCCAAGACATGACACGC 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 LeuserArg----- 235
QY 265 AGGCAATTAAGAGACAGAAAGAGATTGCGTATCTGATCTGAGATCCAGACGTCTC 324
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 -----GlyGlyGlyThrLeuArgAspGlyAlaGlyValIleGlnArgIleuLeu 252
QY 335 TCGTGAATAATAGTATGCGCCACCCCTCAGCAGTGCAGAGCGGTGTGAGCGGT-- 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 LeuserPhemeGlyAlaGluGlnAlaIleProAspProAlaGly-ValIleArgGlyG1 272
QY 383 -----TCTGGAACCACTTCGCGATCACAGAGAACTGCGTGGTCCAGGTGAG 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 yGlyAlaIleGlyLeuThrSerGlyGly-----GlyGlyGlnProGlnTrpG1 288
QY 436 AAAATGCAACAGTACACAGACTACGAAGATCTGAGTTCAGCTGTGTAAGGCATGCG 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 nlyecys-----ArgLeuLeuLeuArgSer-----GluGly----- 298
QY 496 CATCAATGAATAATTTTACCAGCAGACCTATCCAGGAATCGGAATACGACATACGCT 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 -----GluGlyGlyGlyGlySerArg-LeuGluP 308
QY 556 ATTTCAGGATATATGAGGCTCCCTTGTACCTCCCAAGCTTCT----- 605
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 308 hephe-----ValProProlYsAlaSerArgProArgL 319
QY 606 -----GTCTCCACTGAGAGACAAACACAGGG 630
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 euserIleProCysSerThrIleThrAspValArgThrAlaThrAlaLeuGlnMetProA 339
QY 631 ATGTCAGAGTATACACACTGGAAGAGTGCAGAACCTTACC-----TTCAAGGATG 681
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 spArgGluAsnThrPheValValValGluGlyProSerGluTyrIleLeuGlnThr 359
QY 682 TCAGAACCAACCGCTTAAAGGATTCAATACAAAATAAACAAACACCTTTTGGCAC 741
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 hrAspAlaLeuHisValIleValAlaTrp---ValSerAspIleGlnGluCysLeuSerProG 378
QY 742 CTCCTGCGCTGTATCACTCTCCCAAGAAAGTACCAACCTTACCCCGACACACAG 801

```

```

Db 378 lYProCysProAlaIleSer-----ProArgProMetThrLeuPro- 391
QY 802 AGGAGAGCAGTGCATACTTCGCTCCAAAGCCACCTTTCAGAAATCCAGAGGGGCCCA 861
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 -----LeuAlaProGlyThrSerPhe----- 398
QY 862 GGCAGAGGAGTGCAAAAGACTTCAGTAGGGTCTTGAGCAGACAGAAATCTCACACC 921
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 -----LeuThrLysAspAsnThrGlnSerLeuGlnLeuProCysLeuAsnHisSerG 416
QY 922 ACACAAGCCAGAAATCTTCT-----TGCCCATCATCAAAACCAACACAGAGAGCTC 975
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 luserLeuProSerGlnAspLeuLeuGlyProSerG1userAsnAspArgLeuSerG 436
QY 976 CACCTGCCATTGSC-----AGCTCTCTACATGSCAGGAAGACCA 1017
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 lnglyAlaIleTyrGlyIleuserAspArgProSerAlaSerPheSerProSerAlas 456
QY 1018 GTATACAGCCAGACCATACAGGTAGCATGAC-----C 1053
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 456 erIle---AlaAlaSerHisPheAspSerMetGlyLeuLeuProProGluLeuProProA 475
QY 1054 ACTGCTGCTGAGATGCCAGAGCTGCAGCCAGCCACAGCCCTCGAATGCTGCCCTATG 1113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 rgIleProIleGluGlnGlyProProAlaGlyThrValHisProLeuSerThrProTyrP 495
QY 1114 AAAACAACTCGAGGAACCTGACCCCAAGCCTGATGAGAAGATGTGCGCAGA 1173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 roProLeuAspThr-----ProGluAlaIleThrGlySerPheLeuGln 511
QY 1174 ATGAA-----TGACATTTGGAG 1191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 511 lYglnAlaGlyGlyGlyGlyGlyAspGlnProLeuSerGlyTyrProTyrPheHisGly 531
QY 1192 AATACAGTCCGACGAGCAGTGCAGATGTGTATATGAA-----GAGAACAGATGTA 1245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 531 etLeuserArgLeuLysAlaIleGlnLeuValLeuGlnGlyGlyThrSerHisGly 551
QY 1246 CTTTTCGTCGCGAGACTGCTCTACAAATCCAAAGCAGACCAATATGTTGGTGTG 1305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 alPheLeuValArgGlnSerGluThrArg--ArgGlyGly--TyrValLeuThrPhea 569
QY 1306 TTTATGGGACCAAGCTTACATGTGAATAATCGCTTCCTCGAGAGCAATCAACAGTTTG 1365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 snPheGlnGlyAlaIleLysHisLeuArgLeuSerLeuAsnGluGlyGln----- 586
QY 1366 CCTGGGCACAGCACTACAGAGGAATGAGATG--TTGATTCGTGGAAGACATCATTG 1422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 587 -----CysArgValGlnHisLeuThrPheGlnSerIlePheAspMetLeuG 602
QY 1423 AACATACATATATTTCCCATCTGCTATATACATGGG 1460
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 lnhIsPheArgValHisProIleProLeuGlnuserGly 614

```

RESULT 12
 US-08-353-550-8
 Sequence 8, Application us/08353550
 Patent No. 5744313
 GENERAL INFORMATION:
 APPLICANT: Williams, Lewis T.
 APPLICANT: Cavanaugh, William M.
 TITLE OF INVENTION: No. 5744313: Protein Domain Which Binds
 NUMBER OF INVENTION: Tyrosine Phosphorylated Proteins
 CORRESPONDENCE ADDRESSES:
 ADDRESSER: Townsend and Townsend Kourile and Crew
 STREET: One Market Plaza, Steuart Tower, Suite 2000
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94105
 COMPUTER READABLE FORM:

[illegible]

Db 408 OGlyGlnTyrValLeuThr-----G1 415
 QY 1343 CCTGAGAGCAATCAACAGTTTGCCCTG-----GGCAGAGACTACG 1384
 Db 415 yLeuGlnSerGlyGlnProLysHisLeuLeuValAspProGluGlyValArgTTh 435
 QY 1385 AGGAATGAGATGTTTATTCCTGTGGAAGACATCATTTGAACACTAC-----ACATATTT 1438
 Db 435 rLysAspHisArgPheGlnSerValSerHisLeuIleSerTyrHisMetLaspAsnHisLe 455
 QY 1439 TCCATTCTGTATATAGTGAAGAAGAGCTGCAGCAGGAAGAGTGCCTACCTAC 1498
 Db 455 uProIleIle-----SerAlaGlySerIleuLys---LeuG1 467
 QY 1499 CCAGCCACTG 1508
 Db 467 nGlnProVal 470
 RESULT 14
 US-08-426-509A-4
 ; Sequence 4, Application US/08426509A
 ; Patent No. 6326469
 ; GENERAL INFORMATION:
 ; APPLICANT: Ullrich, Axel
 ; APPLICANT: Gishizsky, Mikhail
 ; APPLICANT: Sures, Imman G.
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
 ; TITLE OF INVENTION: TYROSINE KINASES
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York,
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/426,509A
 ; FILING DATE: 21-APR-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/232,545
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7683-0074-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-790-9090
 ; TELEFAX: 212-869-9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 675 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: NO. 6326469e
 ; US-08-426-509A-4
 Alignment Scores:
 Pred. No.: 0.000276 Length: 675
 Score: 133.00 Matches: 62
 Percent Similarity: 39.92% Conservative: 41
 Best Local Similarity: 24.03% Mismatches: 107
 Query Match: 4.35% Indels: 48
 DB: 4 Gaps: 11

US-09-856-061-1 (1-1721) x US-08-426-509A-4 (1-675)
 QY 750 CCTGCTATACACTCTCCCAAGAAATACCAACCTTACCCAGACACAGAGAGAGC 809
 Db 141 ProGlyCysThrLeuThrLeuTrpGluAlaTyrAlaAsnLeuHisThrAlaValaGlnGluLys 160
 QY 810 AGTCATACCTTCGCTCCAAAGCCACTTCCAGAA---GTCCAGAGGGGGCCAGG--- 863
 Db 161 HisArgVal-----ProThrPheProAspArgValLeuLysIleProAlaGala 176
 QY 864 -----CAGAGAGTGCMAAAGACTTCAGTAGGGTCTGTGAGCA----- 902
 Db 177 ValProValLeuLysMetLaspAlaProSerSerSerThrThrLeuAlaGlnTyrAspAsn 196
 QY 903 GAAGAGAAATCTCACACACAGACAGCAACCCAGAAATCTTCTTGCCATCATCAACCAAAAC 962
 Db 197 GluSerLysLysAsnTyrGlySerGlnProProSerSerSerThrSerLeuAlaGlnTyr 216
 QY 963 ACACAGAAAGTCCACCTGCCATTCGCCAGCTCT-----TCTACATGCCA 1007
 Db 217 AspSerAsnSerLysLysIleTyrGlySerGlnProAsnPheAsnMetGlnTyrIlePro 236
 QY 1008 GGA-----AAGCAGATATACAGCCAGA 1031
 Db 237 ArgGlnAspPheProAspTrpTrpGlnValArgLysLeuLysSerSerSerSerGlu 256
 QY 1032 GACCATACAGGTAGCATGACAGCAGCACTGTCTCTCAGAGATCCAGCTGCAGCCAGCCAC 1091
 Db 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
 QY 1092 AGCCCTGGAATGCTGCCCTATGAAACACAACTCGGAGAAACCTGACCCCAAAAGCCT 1151
 Db 271 ThrThrSerLysIleSerTrpGlnPheProGlnSerSerSerGlu----- 286
 QY 1152 GATGAGAGAGATGTCTGCAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1211
 Db 287 GluGlnGlnAsnLeuAspAspTyrAspTrpPheAlaGlnLysIleSerArgSerGlnSer 306
 QY 1212 GAAGATGCTTATGAAAGAGACAGACAGAGTACTTTTGGTCCGAGCTGCTTACA 1271
 Db 307 GluGlnLeuLeuArgGlnGlyGlyGlnGlyGlnGlyGlnGlnGlnGlnGlnGlnGln 323
 QY 1272 AAATCCAGAGCAGAACCATATGTTGGTGGTGGT-----TATGGAGAACAGCTC 1322
 Db 324 SerSerGlnValGlyMetTyrThrValSerLeuPheSerLysAlaValAsnAspLysLys 343
 QY 1323 TACAATGTGAATAATCCGTTTCTCGAGAGCAATGATGATGATGATGATGATGATGATG 1382
 Db 344 GlyThrValLysHisTyrHisValHisThrAsnAlaGlnAsnLysLeu-----TyrLeu 361
 QY 1383 CGAGAAATGAGATGTTGATTCGTGGAAGACATCATTTGAACACTACATAT 1436
 Db 362 AlaGlnAsnTyrCysPheAspSerIleProLysLeuIleHisTyrHisGlnHis 379
 RESULT 15
 PCT-US95-05008-4
 ; Sequence 4, Application PC/TUS9505008
 ; GENERAL INFORMATION:
 ; APPLICANT: Sugen, Inc.
 ; APPLICANT: 515 Galveston Drive
 ; APPLICANT: Redwood City, California 94063-4720
 ; APPLICANT: United States of America
 ; APPLICANT: Wissenschaften E.V.
 ; APPLICANT: Hofgarten Str. 2
 ; APPLICANT: Munchen 80539
 ; APPLICANT: Germany
 ; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
 ; TITLE OF INVENTION: Kinases
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas

Tue Apr 22 16:29:00 2003

us-09-856-061-1.n2p.raii

Page 16

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```

Alignment Scores:	
Pred. No.:	0.000276
Score:	133.00
Percent Similarity:	39.92%
Best Local Similarity:	24.03%
Query Match:	4.35%
DB:	5
Length:	675
Matches:	62
Conservative:	41
Mismatches:	107
Indels:	11
Gaps:	11

QY	750	CGTGTATACATCTTCCCAAGATACCAACCTTACCCCAACACACACAGAGAGAC	809
		:	
Db	141	ProglycylthreuleutprglnalaTyAlaAsnLeuHsthAlaValAsnIugluLys	160
QY	810	AGTCATACACTTCGCTCCAAAGCCCACTTTCAGAA--GTCCAGAGGGGGCCCAAG--	863
Db	161	HisArgVal-----ProthPheProAspArgValLeuLysIleProArgAla	176
QY	864	-----CAGAGAGCTGCAAAAGACTTGAGTAGGGTCTCTTGAGACA-----	902
		:::	
Db	177	ValProValLeuLysMetAspAlaProSerSerThreuleuIaIIntyrAspAsn	196
QY	903	GAGCAGACATCTCACCCACAGCAAAAGCCAGAAATCTTCTTGCCCATCATCAACCAAAAC	962
		:: : :	
Db	197	GluserylLysAsnTyrGlySerGlnProProSerSerThreuleuIaIIntyr	216
QY	963	ACACAGAAAGCTCCACCTGCCATTGCCAGCTC-----TCTCATGCCA	1007
Db	217	AspSerAsnSerLysLysIleTyrGlySerGlnProAsnPheAsnMetCIntIryIlePro	236
QY	1008	GGA-----AAGCAGACAGTAAACCCACA	1033
Db	237	ArgGlnAspPheProAspTyrPrpGlnValArgLysLeuLysSerSerSerSerGlu	256
QY	1032	GACCATACAGCTAGCATGCAGCAGCTGCTCTGCATCAGATGCCAAGTCGACAGCCAC	1091D
Db	257	AspValAlaSerSerAsnGlnLys-----GluArgAsnValAsnHts	270
QY	1092	AGCCCTCGAATGCTGCCCTATGAAAAACACAAACTGGGAAACCTGACCCCAACAAAGCT	1151D

```

Db 271 ThrTrsSerLysIleSerTrpGluPheProGluSerSerSerSerLys----- 286
      ::::: ::::: |||
      ::::: :::::

QY 1152 GATGGAAGCATGTCCTGGCAGCAATGATGTTACATTGGAGCAATACAGTCGCCAGGCACTG 1211
      ::::: ::::: ::::: ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

Db 287 GluGluGluAsnLeuAspAspTyrAspTrpPheAlaGlyAsnIleSerArgSerGlnSer 306
      ::::: ::::: ::::: ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

QY 1212 GAAGCATGTTGTTAAATGAAGAAGCAAGAGATGACTTTTGGTCGAGACATCGCTACA 1271
      ::::: ::::: ::::: ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

Db 307 GluGlnIleuLeuArgGlnLysGlyLysGluGlyAlaPheMetValArgAsn----- 323
      ::::: ::::: ::::: ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

QY 1272 AAATCCAGGCGACCAACCATATGTTTGGTGCTGTTT-----TATGGCAACAGGCTC 1322
      ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

Db 324 SerSerGlnValGlyMetTyrThrValSerLeuPheSerLysAlaValAsnAspLysLys 343
      ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

QY 1323 TACAATGTGMAAAATCCGCTTTCCTCGAGACCAATCAACAGTTGGCCCTGGGCACAGGACTA 1382
      ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

Db 344 GlyThrValLysHisTyrHisValHisThrAspAsnAlaGluAsnLysLeu-----TyrLeu 361
      ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

QY 1383 CGAGGAATGAGATGTTTGATTTCTGGAGAAAGCATGATGCACTACACATAT 1436
      ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

Db 362 AlaGluAsnTrpLysPheAspSerIleProLysLeuIleHisTyrHisGlnHis 379
      ::::: ::::: |||
      ::::: ::::: ::::: ::::: :::::

```

Search completed: April 21, 2003, 12:32:57
Job time : 39.2293 secs

2

GenCore version 5.1.4-P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 12:20:58 ; Search time 43.7798 Seconds
(without alignments)
5944.816 Million cell updates/sec

Title: US-09-856-061-1
Perfect score: 3057
Sequence: 1 acgagggccaactgcaccag.....taaaaaaaaaaaaaaaaa 1721

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 288829 segs, 75613885 residues

Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame-n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09856061/runat_21042003_113021_3809/app_query.fasta.1.3150
-DB=Published.Applications_AA -QMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS-human40.cdt -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09856061 -CCGN_1.1_34 -runat_21042003_113021_3809
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications_AA:*

1: /cgn2.6/prodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2.6/prodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2.6/prodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2.6/prodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2.6/prodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2.6/prodata/2/pubppa/US07_PUBCOMB.pep:*
7: /cgn2.6/prodata/2/pubppa/PCTUS_PUBCOMB.pep:*
8: /cgn2.6/prodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2.6/prodata/2/pubppa/US09_NEW_PUB.pep:*
10: /cgn2.6/prodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2.6/prodata/2/pubppa/US10_NEW_PUB.pep:*
12: /cgn2.6/prodata/2/pubppa/US10_PUBCOMB.pep:*
13: /cgn2.6/prodata/2/pubppa/US60_NEW_PUB.pep:*
14: /cgn2.6/prodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1370.5	44.8	44.3	US-09-966-955A-2
2	1316.5	43.1	428	US-09-966-955A-4
3	1109.5	36.3	353	US-09-966-955A-6
4	212	6.9	46	US-09-864-761-38652

Result	Score	Query Match Length	ID	Description
5	136.5	4.5	370	US-09-764-668-764
6	133	4.4	675	US-10-186-399-3
7	133	4.4	675	US-09-977-260-4
8	133	4.4	675	US-09-977-261-4
9	133	4.4	675	US-09-977-269-4
10	133	4.4	688	US-10-081-980B-1
11	133	4.4	724	US-10-081-980B-3
12	133	4.4	724	US-09-962-929-4
13	130	4.3	724	US-09-962-929-2
14	128.5	4.2	511	US-09-864-761-34590
15	128	4.2	651	US-10-186-399-2
16	128	4.2	1317	US-09-963-896-7
17	125	4.1	524	US-09-764-868-761
18	124.5	4.1	593	US-09-920-021A-3
19	124.5	4.1	597	US-10-038-010-22
20	123.5	4.0	1056	US-10-161-510-10
21	123.5	4.0	1736	US-09-919-97-98
22	122.5	4.0	1163	US-09-932-257A-3
23	121	4.0	1325	US-09-864-761-35612
24	121	4.0	2442	US-10-109-886-10
25	119	3.9	659	US-09-977-260-8
26	119	3.9	659	US-10-045-202-2
27	119	3.9	659	US-09-977-261-8
28	119	3.9	659	US-09-977-269-8
29	118.5	3.9	526	US-09-801-368-362
30	118	3.9	40	US-09-864-761-44023
31	117.5	3.8	522	US-09-764-868-1138
32	117.5	3.8	659	US-08-965-272-2
33	117.5	3.8	945	US-10-121-882-2
34	117.5	3.8	2507	US-09-819-104A-2
35	117.5	3.8	372	US-10-029-180-10
36	117	3.8	543	US-09-977-260-14
37	117	3.8	543	US-09-977-261-14
38	117	3.8	543	US-09-977-269-14
39	117	3.8	543	US-09-801-368-140
40	117	3.8	1302	US-09-977-260-13
41	116.5	3.8	536	US-09-929-260-10
42	116.5	3.8	536	US-09-977-261-13
43	116.5	3.8	536	US-09-977-269-13
44	116.5	3.8	2005	US-09-735-367B-3
45	116.5	3.8	2005	US-09-735-367B-3

ALIGNMENTS

RESULT 1
US-09-966-955A-2
Sequence 2, Application US/09966955A
Patent No. US20020155563A1
GENERAL INFORMATION:
APPLICANT: Perez-Villar, Juan J.
APPLICANT: Chang, Han
APPLICANT: Yang, Wen-Pin
APPLICANT: Wu, Yuli
APPLICANT: Whitney, Gena S.
APPLICANT: Kanner, Steven B.
TITLE OF INVENTION: Identification and Cloning of a Full-length Human
TITLE OF INVENTION: Clink-related Gene, Misp (Mast Cell Immunoreceptor)
TITLE OF INVENTION: Signal Transducer
FILE REFERENCE: 3053-4113US1
CURRENT APPLICATION NUMBER: US/09/966,955A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/270700
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 443
TYPE: PRT
ORGANISM: HUMAN
FEATURES:
OTHER INFORMATION: HUMAN FULL-LENGTH MIST CDNA CLONE #8, TRANSLATED
OTHER INFORMATION: AMINO ACID SEQUENCE

US-09-966-955A-2

Alignment Scores:

Pred. No.:	4,22e-99	Length:	443
Score:	1370.50	Matches:	278
Percent Similarity:	73.30%	Conservative:	46
Best Local Similarity:	62.90%	Mismatches:	113
Query Match:	44.83%	Indels:	5
DB:	9	Gaps:	4

US-09-856-061-1 (1-1721) x US-09-966-955A-2 (1-443)

```

OY 213 GCTAGTTCAGATCCCTCTCTTTCCAGGTCCCAAGGACATGACGAGGCGCAT 272
    |||||
DB 2 AAGGluLeuLysIleProLeuThrArgGlnValProAlaThrMetAsnArgGlnGlyAsn 21
OY 273 AAAGAGCAACAGAGAGATTCGGTATCTGAGATTCAGAACGCTCTCTGCTGANA 332
    |||||
DB 22 ArgLysThrThrLysGlnGlySerAsnAspLeuLysPheGlnAsnPheSerLeuProLys 41
OY 333 AATAGTCTATGGCGCAAGCTCCAGACGTCCCAAGGCGGCTGAGCGGTTCTGAGACCA 392
    |||||
DB 42 AsnArgSerThrProArgLysIleAsnSerAlaThrGlyGlnLysArgMetAsnLysPro 61
OY 393 CTTCGGGATCAGAGAGAACTTGGCTGGGCTCCAGGTGAGAGAAAATGCAACAGTAC 452
    |||||
DB 62 LeuLeuAspTrpGluArgAsnPheAlaValLeuAspGlyAlaLysGlyHisSerAsp 81
OY 453 AACGACTACGAGATCTCTGAGTCCAGCTGCTGAAGGCGCATGCGCATGAAATTTTA 512
    |||||
DB 82 AspAspTrpLysAspProLysLeuArgMetGlnGluTrpGlnSerIleLysIleLeu 101
OY 513 CCAGCGCAGACATCTCAGAGATCCGAAATACAGATACAGCATTCCTCAGATATGATG 572
    |||||
DB 102 ProAlaArgProLysGlnSerGlnLysArgLysAlaAspThrHisLysValAlaMet 121
OY 573 GAGGCTCCGCTCTGTACCTCCAGGCTCTGTCTCAGTACGAGACCAAGGAT 632
    |||||
DB 122 AspThrProLeuProLeuAspThrArgThrSerIleGlyLysProThrTrpHis 141
OY 633 GTGAGATGACACAGCTGGAAGAGTGAAGAGCTTACCTTCAGAGATGAGAGAGCA 692
    |||||
DB 142 ThrGln---ThrArgLeuGlnArgValAspLysProLysSerLysAspValArgSerGln 160
OY 693 CGCTTAAAGATTCGAAATACGAAATAAACAAGACTCTTGGCCACTCTCGGCT 752
    |||||
DB 161 AsnIleLysGlyAspAlaSerValArgLysAsnLysIleProLeuProProArgPro 180
OY 753 GCTATCCTTCGCCCAAGAGTACCAACCTTACCCCGACAGACAGAGAGAGT 812
    |||||
DB 181 LeuIleThrLeuProLysLysLysIleProLeuProPro---GluProGlnSerSerArg 199
OY 813 GCATATCTTGGCTCCAGAGCTTTCAGAGATTCAGAGGCGGCGAGGAGGAGT 872
    |||||
DB 200 ProProLeuSerGlnArgHisThrPheProGlnValGlnAlaMetProSerGlnIleSer 219
OY 873 GCAAAAGCTCAGTAGGCTCTTGGAGAGAGAGAGATTCACACAGACAGACCA 932
    |||||
DB 220 LeuArgAspLeuSerGlnValIleGlnValAlaGlyLysValProHisAsnGlnArgLysPro 239
OY 933 GAATCTTCTTCCCATATCATCAACCAAGAGAGAGAGTCCAGCTGCGATTCGACG 992
    |||||
DB 240 GluSerThrHisLeuLeuGlnAsnGlnHisThrGlnIleLysIleProLysAlaIleSerSer 259
OY 993 TCTTCTACATGCGAGAGAAAGCATATACAGACAGACAGCATACAGTAGTACAG 1052
    |||||
DB 260 SerSerPheThrThrSerAsnHisSerValGlnAsnArgAsnHisArgGlyGlyMetGln 279
OY 1053 CACGTCTCTGCTGAGAGATGCGCAAGCTGACAGCCAGCCAGAGCCCT-----CGAATGCTG 1106
    |||||
DB 280 ProCysSerProGlnArgCysGlnProProLysSerCysSerProHisGlnAsnIleLeu 299
OY 1107 CCTATGAAAAACAACACTCGGAGAAACGTACGCCCAAGAGCTGATGAGAGAGATGTC 1166

```

```

DB 300 ProTyrLysTyrThrSerThrArgProProPheProLysArgSerAspArgLysAspVal 319
OY 1167 TGGCAAGATGATGGTACCTTTGGAGATACAGTCCGACAGGCTGGAAGATGTGTTAATG 1226
    |||||
DB 320 GlnHisAsnGlnTrpTyrLysGlyLysIleGlyLysSerArgGlnAlaValGlnAlaPheMet 339
OY 1227 AAAGAGCAACAGAGATGATCTTTTGGTCCGAGACTGCTGACAAATCCAGGACGAA 1286
    |||||
DB 340 LysGlnAsnLysAspGlySerPheLeuValArgAspCysSerThrLysSerLysGlnGly 359
OY 1287 CCATATGTTTGGTGGTGTGTTTATGGACAAGCTTCAATGTAATAATCCGTTTCTG 1346
    |||||
DB 360 ProTyrValLeuAlaValPheTyrGlnLysValTyrAsnValLysIleArgPheLeu 379
OY 1347 GAGACATCAACACTTTCGCTGGGACAGAGACTACAGAGAAATGACATGTTGATTC 1406
    |||||
DB 380 GluArgAsnGlnGlnPheAlaLeuGlyThrGlyLeuArgGlyAspGlyLysPheAspSer 399
OY 1407 GTGAAAGCATCATGTAACACTACACATATTTTCCATCTGCTAATAGATGGGAAAGAC 1466
    |||||
DB 400 ValGlnAspIleIleGlnHisTyrLysAsnPheProIleIleLeuIleAspGlyLysAsp 419
OY 1467 AAG---CGTCAGCAGAGAGAAACAGTCTACCTACAGCCAGCTGCTGCGAGGCTC 1523
    |||||
DB 420 LysThrGlyValHisArgLysGlnCysHisLeuThrGlnProLeuThrArgHis 439
OY 1524 CTCTCTC 1529
    |||||
DB 440 LeuLeu 441

```

RESULT 2

```

US-09-966-955A-4
: Sequence 4, Application US/09966955A
: Patent No. US2002015563A1
: GENERAL INFORMATION:
: APPLICANT: Perez-Villar, Juan J.
: APPLICANT: Chang, Han
: APPLICANT: Yang, Wen-Pin
: APPLICANT: Wu, Yuli
: APPLICANT: Whitney, Gena S.
: APPLICANT: Kanney, Steven B.
: TITLE OF INVENTION: Identification and Cloning of a Full-length Human
: TITLE OF INVENTION: Clk-related Gene, MIST (mast Cell Immunoreceptor)
: TITLE OF INVENTION: Signal Transducer)
: FILE REFERENCE: 3053-4113US1
: CURRENT APPLICATION NUMBER: US/09/966,955A
: CURRENT FILING DATE: 2001-09-28
: PRIOR APPLICATION NUMBER: 60/237030
: NUMBER OF FILING DATE: 2000-09-29
: NUMBER OF SEQ ID NOS: 52
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4
: LENGTH: 428
: TYPE: PRT
: ORGANISM: HUMAN
: FEATURE:
: OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #7, AMINO ACID
: OTHER INFORMATION: SEQUENCE
US-09-966-955A-4

```

Alignment Scores:

Pred. No.:	7.05e-95	Length:	428
Score:	1316.50	Matches:	266
Percent Similarity:	72.90%	Conservative:	46
Best Local Similarity:	62.15%	Mismatches:	111
Query Match:	43.07%	Indels:	5
DB:	9	Gaps:	4

US-09-856-061-1 (1-1721) x US-09-966-955A-4 (1-428).

```

OY 255 ATGACCAAGGAGGAGATGAAAGAGACGAAAGAGATGCGTGATCTGAGATTTCAG 314
    |||||

```



```

US-09-764-868-764, Application US/09764868
RESULT 5
US-09-764-868-764, Application US/09764868
US-09-856-061-1 (1-1721) x US-09-864-761-38652 (1-46)
QY 1239 GATGCTACTTTTGGTCGCAGCACTCCTCTACAAATCCAGCGAACAACCATATGTGTTG 12988
Db 1 AspliyserPheLeuValAlaArgAspCysSerThrIlySerLySglUclUpotyrvAlleu 20
QY 1299 GTGTCGTTTTATGGGAACAGGCGTCAATGCAATAATCCGTTTCTCGAGACGACATCAA 13588
Db 21 AlavAlaPheIyrlGlusnLysValItyrAsnValLysIIleArpPheLeuCluaRgsmIn 40
QY 1359 CAGTTTGCCCTGGGCACA 1376
Db 41 Glaphealaleuglythr 46
US-09-864-761-38652
Alignment Scores:
Pred. NO.: 1,04e-08 Length: 46
Score: 212.00 Matches: 41
Percent Similarity: 91.30% Conservative: 1
Best Local Similarity: 89.13% Mismatches: 4
Query Match: 6.93% Indels: 0
BB: 10 Gaps: 0
US-09-864-761-38652
OTHER INFORMATION: MAP TO AC005599.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 5.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN SWISSPROT HIT: Q60787, EVALUE 9.00e-10
US-09-864-761-38652
FEATURE:
ORGANISM: Homo sapiens
TYPE: PRT
LENGTH: 46
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 38652

```

```

Patent No. US20020168711A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 764
LENGTH: 370
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (340)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (351)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-764

```

```

US-09-856-061-1 (1-1721) x US-10-186-399-3 (1-675)
OY 750 CCTGATACACTCTCCCAAGAGTACCAACCTTACCCCAAGCACCAGAGAGAGC 809
DB 141 ProGlyCysThrLeuTrpGluAlaIleValAsnIleHisThrAlaValAsnGluIuLys 160
OY 810 AGTCATACCTTCCGTCACAAAGCCCTTCCAGAA--GTCCAGAGGGGGCCAGG--- 863
DB 161 HisArgVal-----ProThrPheProAspArgValLeuIleProArgAla 176
OY 864 -----CAGAGAGTGCACAAAGACTTCAGTAGGGTCTTGAGCA----- 902
DB 177 ValProValLeuLysMetAspAlaProSerSerSerThrThrLeuAlaGlnIleTyrAspAsn 196
OY 903 GAAGAGATCTCACCAACAGCAAGCAAGCAATCTTCTTCCCATCATCAACCAAAAC 962
DB 197 GluSerLysLysAsnIleTyrGlySerGlnProProSerSerSerThrSerLeuAlaGlnIleTyr 216
OY 963 ACACAGAGAGTCCACCTCCATTCGACAGCTCT-----TCTACATGGCA 1007
DB 217 AspSerAsnSerLysLysIleTyrGlySerGlnProAsnPheAsnMetGlnIleTyrIlePro 236
OY 1008 GGA-----AACGACAGATATACACCCAGA 1031
DB 237 ArgGluAspPheProAspTrpTrpGlnValArgLysLeuLysSerSerSerSerGlu 256
OY 1032 GACCATACAGTACGATGACGACACTGCTGCTCAGAGATGCGCAGCCAGCCAC 1091
DB 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
OY 1092 AGCCCTCGAATCTGCTCCCTTAACAAACAACTCGGAGAAACCTGACCCCAAGGCT 1151
DB 271 ThrThrSerLysIleSerTrpGlnPheProGluSerSerSerSerGlu----- 286
OY 1152 GATGAGAGAGTGTGCGCAGAAATGATGATGATGAGAAATACAGTCCGACAGCACTG 1211
DB 287 GluGluGluAsnLeuAspAspTyrAspTrpPheAlaGlyAsnIleSerArgSerGlnSer 306
OY 1212 GAAGATGTGTTAAATGAAGAGCAAGATGATCTTTTGGTCCGAGAGTCTCTACA 1271
DB 307 GluGlnLeuLeuArgGlnLysGlyGlyAlaPheMetValArgAsn----- 323
OY 1272 AAATCCAGGCGAAGCAATATGTTTGGTGGTCTTT-----TATGGAGCAAGGCTC 1322
DB 324 SerSerGlnValGlyMetLysTrpThrValSerLeuPheSerLysAlaValAsnAspLysLys 343
OY 1323 TACAAATGGAATCCGTTCTCGAGAGCAATCAACAGTTTGGCCCTGGGACAGAGCTA 1382
DB 344 GlyThrValLysHisIleTyrHisValHisThrAsnAlaGluAsnLysLeu-----TyrLeu 361
OY 1383 CGAGCAATGAGATGTTGATCTGTGGAGACATCATTTGAAACATACATATAT 1436
DB 362 AlaGluAsnTyrCysPheAspSerLysLeuLysLeuIleHisIleTyrHisGlnHis 379

RESULT 7
US-09-977-260-4
; Sequence 4, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHITZKY, MIKHAIL
; APPLICANT: SURES, ILMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 675
; TYPE: PRT

```

```

; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 2
US-09-977-260-4

Alignment Scores:
Pred. No.: 0.0368 Length: 675
Score: 133.00 Matches: 62
Percent Similarity: 39.92% Conservative: 41
Best Local Similarity: 24.03% Mismatches: 107
Query Match: 4.35% Indels: 48
Dbs: 9 Gaps: 11

US-09-856-061-1 (1-1721) x US-09-977-260-4 (1-675)
OY 750 CCTGATACACTCTCCCAAGAGTACCAACCTTACCCCAAGCACCAGAGAGAGC 809
DB 141 ProGlyCysThrLeuTrpGluAlaIleValAsnIleHisThrAlaValAsnGluIuLys 160
OY 810 AGTCATACCTTCCGTCACAAAGCCCTTCCAGAA--GTCCAGAGGGGGCCAGG--- 863
DB 161 HisArgVal-----ProThrPheProAspArgValLeuIleProArgAla 176
OY 864 -----CAGAGAGTGCACAAAGACTTCAGTAGGGTCTTGAGCA----- 902
DB 177 ValProValLeuLysMetAspAlaProSerSerSerThrThrLeuAlaGlnIleTyrAspAsn 196
OY 903 GAAGAGATCTCACCAACAGCAAGCAAGCAATCTTCTTCCCATCATCAACCAAAAC 962
DB 197 GluSerLysLysAsnIleTyrGlySerGlnProProSerSerSerThrSerLeuAlaGlnIleTyr 216
OY 963 ACACAGAGAGTCCACCTCCATTCGACAGCTCT-----TCTACATGGCA 1007
DB 217 AspSerAsnSerLysLysIleTyrGlySerGlnProAsnPheAsnMetGlnIleTyrIlePro 236
OY 1008 GGA-----AACGACAGATATACACCCAGA 1031
DB 237 ArgGluAspPheProAspTrpTrpGlnValArgLysLeuLysSerSerSerSerGlu 256
OY 1032 GACCATACAGTACGATGACGACACTGCTGCTCAGAGATGCGCAGCCAGCCAC 1091
DB 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
OY 1092 AGCCCTCGAATCTGCTCCCTTAACAAACAACTCGGAGAAACCTGACCCCAAGGCT 1151
DB 271 ThrThrSerLysIleSerTrpGlnPheProGluSerSerSerSerGlu----- 286
OY 1152 GATGAGAGAGTGTGCGCAGAAATGATGATGATGAGAAATACAGTCCGACAGCACTG 1211
DB 287 GluGluGluAsnLeuAspAspTyrAspTrpPheAlaGlyAsnIleSerArgSerGlnSer 306
OY 1212 GAAGATGTGTTAAATGAAGAGCAAGATGATCTTTTGGTCCGAGAGTCTCTACA 1271
DB 307 GluGlnLeuLeuArgGlnLysGlyGlyAlaPheMetValArgAsn----- 323
OY 1272 AAATCCAGGCGAAGCAATATGTTTGGTGGTCTTT-----TATGGAGCAAGGCTC 1322
DB 324 SerSerGlnValGlyMetLysTrpThrValSerLeuPheSerLysAlaValAsnAspLysLys 343
OY 1323 TACAAATGGAATCCGTTCTCGAGAGCAATCAACAGTTTGGCCCTGGGACAGAGCTA 1382
DB 344 GlyThrValLysHisIleTyrHisValHisThrAsnAlaGluAsnLysLeu-----TyrLeu 361
OY 1383 CGAGCAATGAGATGTTGATCTGTGGAGACATCATTTGAAACATACATATAT 1436
DB 362 AlaGluAsnTyrCysPheAspSerLysLeuLysLeuIleHisIleTyrHisGlnHis 379

RESULT 8
US-09-977-261-4
; Sequence 4, Application US/09977261
; Publication No. US20030054527A1
; GENERAL INFORMATION:

```

APPLICANT: ULLRICH, AXEL
 APPLICANT: GISHIZKY, MIKHAIL
 APPLICANT: SURES, IRMINGARD
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
 FILE REFERENCE: 038602/1259
 CURRENT APPLICATION NUMBER: US/09/977,261
 CURRENT FILING DATE: 2001-10-16
 PRIOR APPLICATION NUMBER: 08/232,545
 PRIOR FILING DATE: 1994-04-22
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 4
 LENGTH: 675
 TYPE: PRT
 ORGANISM: Unknown Organism
 FEATURE:
 OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
 OTHER INFORMATION: kinase 2
 US-09-977-261-4

Alignment Scores:
 Pred. No.: 0.0368 Length: 675
 Score: 133.00 Matches: 62
 Percent Similarity: 39.92% Conservative: 41
 Best Local Similarity: 24.03% Mismatches: 107
 Query Match: 4.35% Indels: 48
 Gaps: 11

US-09-856-061-1 (1-1721) x US-09-977-261-4 (1-675)
 QY 750 CCTGCTATCTCTCCCAAGAGTACCAACCTTACCCAGCAGCAGAGAGAGC 809
 DB 141 ProGlycylserThrLeuTrpGluAlaThrAlaAsnLeuHisThrAlaValAsnGluGluLys 160
 QY 810 AGTGCATCTTCGCTCCCAAGCCACCTTCCAGAA--GTCCAGAGGGGGCCAGG-- 863
 DB 161 HisArgVal-----ProThrPheProAspArgValLeuLysIleProArgAla 176
 QY 864 -----CAGAGAGTCAAAAGACTTCAGTAGGGTCCCTGGAGCA----- 902
 DB 177 ValProValLeuLysMetAspAlaProSerSerSerThrThrLeuAlaGlnTyrAspAsn 196
 QY 903 GAAGAGAATCTCCACCAAGCAAGCAAGCAAGATCTTCTGCCATCATCAACCAAAAC 962
 DB 197 GluSerLysLysAsnTyrGlySerGlnProProSerSerSerThrSerLeuAlaGlnTyr 216
 QY 963 ACACAGAAGAGTCCACCTGCCATTGCCAGCTCT-----TCTACATGCCA 1007
 DB 217 AspSerAsnSerLysLysIleTyrGlySerGlnProAsnPheAsnMetGlnTyrIlePro 236
 QY 1008 GGA-----AGCACAGTATACAGCCAGA 1031
 DB 237 ArgGluAspPheProAspTyrTrpGlnValArgLysLeuLysSerSerSerSerSerGlu 256
 QY 1032 GACCATACAGGTAGCATGACACTGTCTGCTCAGAGATGCCAGTCCAGCCAGCCAC 1091
 DB 257 AspAlaIaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
 QY 1092 AGCCCTCAATGTCGCCCTATGAAACAAACAACCTGCGAAGAACCTGACCCCAAAAGCT 1151
 DB 271 ThrThrSerLysLysIleSerTrpGlnPheProGluSerSerSerSerGlu----- 286
 QY 1152 GATGAGAGAGATGTCTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1211
 DB 287 GluGluGluAsnLeuAspAspTyrAspTrpPheAlaGlnLysAsnIleSerArgSerGlnSer 306
 QY 1212 GAAGATGCTTATGAAGAAGCAAGAGATGCTTTTGGTCCGAGACTGCTCTCA 1271
 DB 307 GluGlnLeuLeuArgGlnLysGlySerGlnLysAlaPheMetValArgAsn----- 323
 QY 1272 AATTCACAGGAGCAACCATATGTTTGGGTGGTGT-----TATGGCAACAGGTC 1322
 DB 324 SerSerGlnValGlyMetTyrThrValSerLeuPheSerLysAlaValAsnAspLysLys 343

QY 1323 TACATGTGAATACCCGTTCTCTGAGAGCAATCAAGTTCGCCGTGGCAGAGACTA 1382
 DB 344 GlyThrValLysHisTyrHisValHisThrAsnIaGluAsnLysLeu-----TyrLeu 361
 QY 1383 CGAGAAATGAGATGTTCATTCGTGCGAAGACATCATTCATGAACATCAATAT 1436
 DB 362 AlaGluAsnTyrGlyCysPheAspSerIleProLysLeuIleHisTyrHisGlnHis 379

RESULT 9
 US-09-977-269-4
 Sequence 4, Application US/09977269
 Patent No. US20020082037A1
 GENERAL INFORMATION:
 APPLICANT: ULLRICH, AXEL
 APPLICANT: GISHIZKY, MIKHAIL
 APPLICANT: SURES, IRMINGARD
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
 FILE REFERENCE: 038602/1260
 CURRENT APPLICATION NUMBER: US/09/977,269
 CURRENT FILING DATE: 2001-10-16
 PRIOR APPLICATION NUMBER: 08/232,545
 PRIOR FILING DATE: 1994-04-22
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 4
 LENGTH: 675
 TYPE: PRT
 ORGANISM: Unknown Organism
 FEATURE:
 OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
 OTHER INFORMATION: kinase 2
 US-09-977-269-4

Alignment Scores:
 Pred. No.: 0.0368 Length: 675
 Score: 133.00 Matches: 62
 Percent Similarity: 39.92% Conservative: 41
 Best Local Similarity: 24.03% Mismatches: 107
 Query Match: 4.35% Indels: 48
 Gaps: 11

US-09-856-061-1 (1-1721) x US-09-977-269-4 (1-675)
 QY 750 CCTGCTATCTCTCCCAAGAGTACCAACCTTACCCAGCAGCAGAGAGC 809
 DB 141 ProGlycylserThrLeuTrpGluAlaThrAlaAsnLeuHisThrAlaValAsnGluGluLys 160
 QY 810 AGTGCATCTTCGCTCCCAAGCCACCTTCCAGAA--GTCCAGAGGGGGCCAGG-- 863
 DB 161 HisArgVal-----ProThrPheProAspArgValLeuLysIleProArgAla 176
 QY 864 -----CAGAGAGTCAAAAGACTTCAGTAGGGTCCCTGGAGCA----- 902
 DB 177 ValProValLeuLysMetAspAlaProSerSerSerThrThrLeuAlaGlnTyrAspAsn 196
 QY 903 GAAGAGAATCTCCACCAAGCAAGCAAGATCTTCTGCCATCATCAACCAAAAC 962
 DB 197 GluSerLysLysAsnTyrGlySerGlnProProSerSerSerThrSerLeuAlaGlnTyr 216
 QY 963 ACACAGAAGAGTCCACCTGCCATTGCCAGCTCT-----TCTACATGCCA 1007
 DB 217 AspSerAsnSerLysLysIleTyrGlySerGlnProAsnPheAsnMetGlnTyrIlePro 236
 QY 1008 GGA-----AGCACAGTATACAGCCAGA 1031
 DB 237 ArgGluAspPheProAspTyrTrpGlnValArgLysLeuLysSerSerSerSerSerGlu 256
 QY 1032 GACCATACAGGTAGCATGACACTGTCTGCTCAGAGATGCCAGTCCAGCCAGCCAC 1091
 DB 257 AspAlaIaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
 QY 1092 AGCCCTCAATGTCGCCCTATGAAACAAACAACCTGCGAAGAACCTGACCCCAAAAGCT 1151


```

Db 271 ThrThrserysileserrtrglupheProglUserSerSerSerSerGlu----- 286
QY 1152 GATGACAGAGATGCTGCGCAGCAATGTAATGTACATTGGAGATACAGTCCCGCAGCAGTg 1211
Db 287 GUGUGUUAUnleuAsnLeuAspArgTyrAspTrpPheAlaGlyAsnIleSerArgSerGlnSer 306
QY 1212 GAAGATGCTGTAATGAAGAACAGAGATGTAATGTTTGGTCCGAGACCTGCTTACA 1271
Db 307 GUGUUnleuArgGlnLysGlyLysGlyAlaPheMetValArgAsn----- 323
QY 1272 AATCCAGCGACACCATATGTTTGGTGGTGT-----TATGGAAACAGATC 1322
Db 324 SerSerGlnValGlyMetLysTrpValSerLeuPheSerLysAlaValAsnAspLysLys 343
QY 1323 TACAAATGAAAATCCGTTTCCCGCAGACCAATCAACAGATTGGCCCTGGCGCAGACTA 1382
Db 344 GLyThValLysHisLysTrpHisValHisTrpAsnAlaGluAsnLysLeu-----TyrLeu 361
QY 1383 CCAGGAATGAGATGTTGATTCGTGTGGAAGACATCATGGAACACTACATAT 1436
Db 362 AlAGUAAsnTyrCysPheAspSerIleProLysLeuIleHisLysTrpHisGlnHis 379

RESULT 10
US-10-081-980B-1
; Sequence 1, Application US/10081980B
; Publication No. US2003004137A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Gibbs, E. Michael
; APPLICANT: McNeilsh, John D.
; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form Of
; FILE REFERENCE: PC10206AGPR
; CURRENT APPLICATION NUMBER: US/10/081.980B
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/270,014
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 688
; TYPE: PRT
; ORGANISM: mus musculus
US-10-081-980B-1

Alignment Scores:
Pred. No.: 0.037 Length: 688
Score: 133.00 Matches: 56
Percent Similarity: 46.86% Conservative: 26
Best Local Similarity: 32.00% Mismatches: 75
Query Match: 4.35% Indels: 19
Gaps: 9

US-09-856-061-1 (1-1721) x US-10-081-980B-1 (1-688)
QY 913 CTCACACGACGACAAAGCCGAATCTTCTT---GCCCATCATCAACCAACACACAGAGA 969
Db 250 LeuSerGlnAlaSerSerLysAsnLeuAsnAlaArgValLeuSerGluIlePheSer 269
QY 970 AGAGTCACCTGCGCATTCCTCTCAATGCGCAGAAAGCAAGATATCAAGCA 1029
Db 270 ProValLeuPheArgPheProAla-----AlaSerSerAspAsnTrpGluHisLeuIle 287
QY 1030 GAGACCATACAGATGATGACGACACTGCTCTCGACAGATGCCAAGCTGCAGCCAGCC 1089
Db 288 LysAlaIleGluIleLeuIleSerThr-----GluTrpAsnGlnArgGlnProAla 304
QY 1090 ACAGCCCTCGAATGCTGCCCTATGAACAAACAACTCGGAGAACTGCAGCCCAAAAGC 1149
Db 305 ProAlaLeuProProLysProProLysProThrThrValAlaAsn-----AsnSer 321
QY 1150 CTGATGAGAAGATGCTGTG---CAGAAATGAATGTGATATTGGAATATACAGTCCGAG 1205

```

```

Db 322 MetAsnAsnAsnMetSerLeuGlnAspAlaGluTrpTyrTrpGlyAspIleSerArgGlu 341
QY 1206 GCAGTCGAGATGCTGTTAATGAAGAACAGAGATGTAATGTTTGGTCCGAGACTGC 1265
Db 342 GluValAsnGlnLysLeu-----ArgAspThrAlaAspGlyThrPheLeuValArgAspAla 360
QY 1266 TCTACAAAATCCAGCGACGACCAACCATATGTTTGGTGGTGTATTATGAGAACAGCTAC 1325
Db 361 SerThrLysMetHisGlyAsp---TyrThrLeuThrLeuArgGlyGlyAsn----- 377
QY 1326 AATGTGAAAATCCGTTTCCCTCGAGAGCAATCAACAGTTTGGCCCTGGCGCAGACTACGA 1385
Db 378 AsnLysLeuIleLysIlePheHisArgAspGlyLysTyrGlyPheSerAspProLeu--- 396
QY 1386 GCAATGAGATGTTGATTCGTGTGGAACACATCATGGAACCTGAC 1430
Db 397 -----ThrPheAsnSerValValGluLeuIleAsnHisLysr 408

RESULT 11
US-10-081-980B-3
; Sequence 3, Application US/10081980B
; Publication No. US2003004137A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Gibbs, E. Michael
; APPLICANT: McNeilsh, John D.
; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form
; FILE REFERENCE: PC10206AGPR
; CURRENT APPLICATION NUMBER: US/10/081.980B
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/270,014
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 724
; TYPE: PRT
; ORGANISM: mus musculus
US-10-081-980B-3

Alignment Scores:
Pred. No.: 0.0375 Length: 724
Score: 133.00 Matches: 56
Percent Similarity: 46.86% Conservative: 26
Best Local Similarity: 32.00% Mismatches: 75
Query Match: 4.35% Indels: 19
Gaps: 9

US-09-856-061-1 (1-1721) x US-10-081-980B-3 (1-724)
QY 913 CTCACACGACGACAAAGCCGAATCTTCTT---GCCCATCATCAACCAACACACAGAGA 969
Db 250 LeuSerGlnAlaSerSerLysAsnLeuAsnAlaArgValLeuSerGluIlePheSer 269
QY 970 AGAGTCACCTGCGCATTCCTCTCAATGCGCAGAAAGCAAGATATCAAGCA 1029
Db 270 ProValLeuPheArgPheProAla-----AlaSerSerAspAsnTrpGluHisLeuIle 287
QY 1030 GAGACCATACAGATGATGACGACACTGCTCTCGACAGATGCCAAGCTGCAGCCAGCC 1089
Db 288 LysAlaIleGluIleLeuIleSerThr-----GluTrpAsnGlnArgGlnProAla 304
QY 1090 ACAGCCCTCGAATGCTGCCCTATGAACAAACAACTCGGAGAACTGCAGCCCAAAAGC 1149
Db 305 ProAlaLeuProProLysProProLysProThrThrValAlaAsn-----AsnSer 321
QY 1150 CTGATGAGAAGATGCTGTG---CAGAAATGAATGTGATATTGGAATATACAGTCCGAG 1205
Db 322 MetAsnAsnAsnMetSerLeuGlnAspAlaGluTrpTyrTrpGlyAspIleSerArgGlu 341
QY 1206 GCAGTCGAGATGCTGTTAATGAAGAACAGAGATGTAATGTTTGGTCCGAGACTGC 1265

```



```

Db      342  GLUValAsnGluLysLeu---ArgAspThrAlaAspGlyThrPheLeuValAlaGAspAla 360
OY      1266 TCACAAATTCACAGGACAGACATATGTTGGTGTGTTTATGGAACAAGTCTAC 1325
Db      361  SerThrLysMetHisGlyAsp---TyrThrLeuThrLeuArgLysGlyAsn----- 377
OY      1326 AATGTGAAATCCGTTTCTCGAGACCAATCAACAGTTTCCCTGGGACAGAGACTACGA 1385
Db      378  AsnLysLeuLeuLysLysPheHisArgAspGlyLysTyrGlyPheSerAspProLeu--- 396
OY      1386 GGAATGAGATGTTGTCTGTCTGTGGAAGACATCATGACATGACATAC 1430
Db      397  -----ThrPheAsnSerValValGluLeuLeuAsnHisTyr 408

RESULT 12
US-09-962-929-4
; Sequence 4, Application US/09962929
; Patent No. US20020115058A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Diseases Associated with A
; TITLE OF INVENTION: Expression of PIK3r1
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/668,644
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-962-929-4

Alignment Scores:
Pred. No.: 0.0376 Length: 724
Score: 133.00 Matches: 96
Percent Similarity: 37.32% Conservative: 57
Best local Similarity: 23.41% Mismatches: 169
Query Match: 4.35% Indels: 89
Gaps: 16

US-09-856-061-1 (1-1721) x US-09-962-929-4 (1-724)
OY      371  GTGTGACGCGGTCT------GGAACCACTTCGGATGCACAG-----AAG 409
Db      32  ValAsnLysGlySerLeuValAlaLeuLysPheSerAspGlyGluAlaArgProGlu 51
OY      410  GAAGTGGCTGGGGTCCAGGT-----GGAAGAAATTCACACAGTACAA 454
Db      52  GluIleGlyTyr-LeuAsnGlyTyrAsnGluThrThrGlyGluArgGlyAspPheProGlu 71
OY      455  CGACATCAAGATCTCTGAGTCCAGCTGCTGTAAGGCATGGCCATCAAGAAATTTTAC 514
Db      71  YThrYrValGluTyrIleGlyArgLysLysLysSerProPheProThrProLysProArgPr 91
OY      515  AGCCAGACCTTATCCAGATCGAATCGAATACAGCTATTTCCAGATATGATGA 574
Db      91  OProArgProLeuProValAlaPro---GlySerSerLysThrGluAlaAspValGluGlu 110
OY      575  GGCTCCCTCTGTACTCTCCAGGCTTCT----- 605
Db      110  nGluAlaLeuThrLeuProAspLeuAlaGluInPheAlaProAspIleAlaProPr 130
OY      606  -----CT 607
Db      130  oLeuLeuIleLysLeuValGluAlaIleGluLysLysGlyLeuGluCysSerThrLeuTyr 150
OY      608  CTCACACTGAGAGACAACACGAGGATGTGAGATGACACAGCTGGAAGAGTGCACAGCC 667

```

```

Db      150  rArgThrGlnSerSerSerAsnLeuAlaGluLeuArgGlnLeuLeuAspCysAspThrPr 170
OY      668  TACCTTCAG-----GATCTCAAGAACCAACGCTTTAAAGATTCAA 709
Db      170  oSerValAspLeuGluMetIleAspValHisValIleuAlaAspAlaPheLysArgTyrLe 190
OY      710  ATACACAAATAATACAAACACTCTTGGCCACTCTCGGCTGCTATCACTCTCCCA 769
Db      190  u-----LeuAspLeuProAsnProValIleProAlaAlaVala 202
OY      770  GAAGTCAACACCTTACCCCAACACACAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 827
Db      202  IYrSerGluMetIleSerLeuAlaProGluValGlnSerSerGluTyrIleGluLe 222
OY      828  -----AAGCCACCTTTCAGAAATTCACAGAGGGGCCACAGCAGAG 868
Db      222  uLeuLysLysLeuIleArgSerProSerIleProHisGlnTyrTyrPheuThrLeuGlnTyr 242
OY      869  GAGTCGAAAGACTTCAGTAGGCTCCTTGAGCAGAAAGAAATCTCACACACAGACAAA 928
Db      242  rLeuLeuLysHisPhePheLysLeu-----SerGlnThrSe 254
OY      929  GCCAGATCTTCTGGCCATCATCAACCAACCAACACAGAGAGTCCACTGCTTC 988
Db      254  rSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePheSerProMetLeuPheAr 274
OY      989  CAGCTCTCTCCATATGCCAGAGAAACAGCATATACAGCAGAGACCATATACAGTACAT 1048
Db      274  gPheSerAla-AlaSerSerAspAsnThrGluAsnLeuIleLysValIleGluIleLeuI 294
OY      1049  GCAGCACTGTCTGCTCAGATGCAGATGCAGCGCAGCCAGCCCTCGAATGCTGCC 1108
Db      294  lSerThr-----GluTyrAsnGluArgGlnProAlaProAlaLeuProLysPr 311
OY      1109  CTATGAAACCAACACTCGGAGAAACCTGACCCACAAACCTGTAGAGAGAGATGCTG 1168
Db      311  rProLysProThrThrValAlaAsn-----AsnGlyMetAsnAsnMetSerL 328
OY      1169  G-CAGAT-----GATGTGATATGAGATATACAGTCCGACGAGCAGGAGATGTGTTAA 1224
Db      328  eGlnAsnAlaGluTyrTyrTyrPheLysPheSerAspGlyGluValAsnGluLysLeu- 347
OY      1225  TGAAGAGAACACAGATGATGATTTTGTGCTCCAGACTCTCTACAAATCCAAAGCAG 1284
Db      348  --ArgAspThrAlaAspGlyThrPheLeuValAlaArgAspAlaSerThrLysMetHisGly 367
OY      1285  AACCATATGTTTGTGTGTGTTTATGGAACAAGTCTACAAATCGAATCGCTTCC 1344
Db      367  sp---TyrThrLeuThrLeuArgLysGlyAsn-----AsnLysLeuIleLysLysP 384
OY      1345  TCGAGAGCAATCAACCTGCTCCCTGGCAGAGACTACGAGCAATGAGATGTTGATT 1404
Db      384  heHisArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheSer 400
OY      1405  CTGTGAGAGACATCATGACACTAC 1430
Db      400  eValValGluLeuIleAsnHisTyr 408

RESULT 13
US-09-962-929-2
; Sequence 2, Application US/09962929
; Patent No. US20020115058A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Diseases Associated wit
; TITLE OF INVENTION: Expression of PIK3r1
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; PRIOR FILING DATE: 2001-09-24

```

```

1  PRIOR APPLICATION NUMBER: US 09/632,366
2  PRIOR FILING DATE: 2000-08-03
3  PRIOR APPLICATION NUMBER: GB 24263.6
4  PRIOR FILING DATE: 2000-10-04
5  PRIOR APPLICATION NUMBER: US 60/236,359
6  PRIOR FILING DATE: 2000-09-27
7  PRIOR APPLICATION NUMBER: PCT/US01/00666
8  PRIOR FILING DATE: 2001-01-30
9  PRIOR APPLICATION NUMBER: PCT/US01/00667
10 PRIOR FILING DATE: 2001-01-30
11 PRIOR APPLICATION NUMBER: PCT/US01/00664
12 PRIOR FILING DATE: 2001-01-30
13 PRIOR APPLICATION NUMBER: PCT/US01/00669
14 PRIOR FILING DATE: 2001-01-30
15 PRIOR APPLICATION NUMBER: PCT/US01/00665
16 PRIOR FILING DATE: 2001-01-30
17 PRIOR APPLICATION NUMBER: PCT/US01/00668
18 PRIOR FILING DATE: 2001-01-30
19 PRIOR APPLICATION NUMBER: PCT/US01/00663
20 PRIOR FILING DATE: 2001-01-30
21 PRIOR APPLICATION NUMBER: PCT/US01/00662
22 PRIOR FILING DATE: 2001-01-30
23 PRIOR APPLICATION NUMBER: PCT/US01/00661
24 PRIOR FILING DATE: 2001-01-30
25 PRIOR APPLICATION NUMBER: PCT/US01/00670
26 PRIOR FILING DATE: 2001-01-30
27 PRIOR APPLICATION NUMBER: US 60/234,687
28 PRIOR FILING DATE: 2000-09-21
29 PRIOR APPLICATION NUMBER: US 09/608,408
30 PRIOR FILING DATE: 2000-06-30
31 PRIOR APPLICATION NUMBER: US 09/774,203
32 PRIOR FILING DATE: 2001-01-29
33 NUMBER OF SEQ ID NOS: 49117
34 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
35 SEQ ID NO 34590
36 LENGTH: 511
37 TYPE: PRT
38 ORGANISM: Homo sapiens
39
40 OTHER INFORMATION: MAP TO AB023049.1
41 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.7
42 OTHER INFORMATION: EXPRESSED IN H1100, SIGNAL = 9.3
43 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
44 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
45 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
46 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.2
47 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.8
48 OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.8
49 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
50 OTHER INFORMATION: SWISSPROT HIT: Q02817, EVALUE 3.00e-24
51 OTHER INFORMATION: EST_HUMAN HIT: AA487852.1, EVALUE 2.00e-73
52 US-09-866-761-34590
53
54 Alignment Scores:
55 Pred. No.: 0.0759 Length: 511
56 Score: 128.50 Matches: 104
57 Percent Similarity: 33.04% Conservative: 45
58 Best Local Similarity: 23.06% Mismatch: 163
59 Query Match: 4.20% Indels: 139
60 DB: 10 Gaps: 23
61
62 US-09-856-061-1 (1-1721) x US-09-864-761-34590 (1-511)
63
64 Oy 456 GACATCGAAGATCCTGCAATTCACAGCTGCTGAGGCAATGAGCAATGAAATTTTACCA 515
65 ||| ||| |||::: ||| ||| |||
66 Db 2 ASpgAGnGlnSerProAspAlaCysLeuProProThrVal|ProGluAlaProAlaProPro 21
67 ::|||:::
68 Oy 516 GCCAGCATATTCACGGAATGGAATATGAGCAATACAGCATATTTCCAGGATATGATGAGAG 575
69 ::|||:::
70 Db 22 GlnlypProLeuAanSerGlnSerGln-----LysShiSleuAlaPro 35
71 576 GCTCCCTCTTGTATTACCTCCCAAGGCTTCTGTCTCC---ACTGAGAGACAACACGAGAT 632

```

Db	36	proProLeuLeuSerProLeuLeuProSerIleLysProThrValArgLysThrArgGln	55
Qy	633	GTGAGGATGACACAGCTGGAGAGAGTG-----	GAC 662
Db	56	AspGlySerGlnGluLapProGluLapProLeuSerSerGluLeuGluProPheHisPro	75
Qy	663	AACCCATACCTTCAGAGGATGTCAAGGACCCAGCCGTTTAAAGGATGCAATATCAAAAATA	722
Db	76	LysProLysLleArgThrArgLysSerSerArgMetThrProPheProIaThr-----	93
Qy	723	AACAAGACTCTTTGGCCACCTCTCGGCCCTGTATCACTCTCCCAAGAAAGTACCAACC	782
Db	94	-----SerAlaIaIaProGluProHisProSerThrSerThrAla-----GlnPro	108
Qy	783	TTTACCCCAAGCACACCACCAAGAGAGAC-----	AGTCGATACTTC 821
Db	109	ValThrProLysProThrSerGlnAlaThrArgSerArgThrAsnArgSerSerValLys	128
Qy	822	GGTCCAAAGCCCACTTTTCA-----	842
Db	129	ThrProGluProValValProThrIaIaProGluLeuGlnProSerThrSerThrAspGln	148
Qy	843	-----	GAATCCAGAGGGGGCCAGCAGAGAGTGCA--- 875
Db	149	ProValThrSerGluProThrSerGlnValThrArgLysArgLysSerArgSerVal	168
Qy	876	AAAGACTTCAGTAGGCTCTTGAGCAGAAAGAAAGAAATCTCAACCAACCAACA-----	926
Db	169	LysThrProGluThrValValProThrIaIaLeuGluLeuGlnProSerThrSerThrAsp	188
Qy	927	AAGCAAGATCTTCTTGCCCATCATCA-----	AACCAAAACAGA 965
Db	189	ArgProValThrSerGluProThrSerGlnAlaThrArgGlyArgLysAsnArgSerSer	208
Qy	966	CAGAAGAGTCCA---CCTGCAATTCGCCAGCTCTCTCATATCCAGGAAGACACA-----	1017
Db	209	ValLysThrProGluProValVal---ProThrIaIaProLeuGlnProSerThrSerThr	228
Qy	1018	-----GTATACAAAGCCAGAGACCAATACAGTAGATGCACACACTGTCCTGCTCAG	1067
Db	228	rAspGlnProValThrSerGluProThrArgGlnAlaThrArgGlyArgLysAsnArgS	248
Qy	1068	AGATGCCAAGCTG---CAGCCAGCC-----ACAGCCCTCGAATGCTGCCCTATGAAGAAC	1118
Db	248	rSerValLysThrProGluProValValProThrIaIaProGluLeuArgPro---SerThr	267
Qy	1119	ACAACATCGGAGAAACCTGACCCCAACAGCCTGATGAGACAGAGATGCTGCACAGATGAA	1178
Db	267	rSerThrAspArgProValThrProLysProThrSerArgThrThrArgSerArgThrAs	287
Qy	1179	TGTGATCANTGGAATACA-----	1197
Db	287	metSerSerValLysThrProGluThrValValProThrIaIaProGluLeuGlnIleS	307
Qy	1198	-----	GTGCCAGAGCAGGTGAGAGATGTGTTA 1223
Db	307	rThSerThrAspGlnProValThrProLysProThrSerArgThrThrArg-----	324
Qy	1224	ATGAAGAAGAACAGATGCTACTTTTGGTCCGAGACTGCTCTACAAAATCCAAAGCA	1288
Db	325	-----SerArgThrAsnMet-----	SerSerValLysAsnProGluS 337
Qy	1284	GAACCATATGTTTGGTGGTGTTTATGAGAACAGAGTCAACATGTGAAATCCGTTTC	1343
Db	337	r-ThrValProIleAlaProGluLeu---ProProSerThrSerThrGluGlnProValT	356
Qy	1344	CTCGAGAGCAATCAACAGTTTCCCTTGCGGACAGGACTACAGAGAAATGAGATGTTGAT	1403
Db	356	hrProGlu-----	ProThrSerArgIaThrArg----- 365
Qy	1404	TCTGTGGAAGACATCTTGAACACTACACATATTTTCCATTCCTGTGTAATAGATGGGAA	1466

QY	1464	GACAGAGCTGCACGACGAGAAACAGTGTACTACACCCAGCCACTGCCTTCGCCAGGCTC	1523
Db	371	erserGlySerThrProGluThrLeuValPro-----ThrAlaProLysLeuGlu	388
QY	1524	CTTCTCAGTCAGTACCTCCGACGAGGACT	1552
Db	388	roserThrSerThrAspGlnProValThr	397
RESULT 15			
US-10-186-399-2			
Sequence 2, Application US/10186399			
Patent No. US20020173481A1			
GENERAL INFORMATION:			
APPLICANT: Ekman, Niklas			
APPLICANT: Arighi, Elena			
APPLICANT: Vastrik, Imre			
APPLICANT: Tamagnone, Luca			
APPLICANT: Altalo, Kari			
TITLE OF INVENTION: REGULATION OF VASCULAR ENDOTHELIUM USING BMX TYROSINE			
TITLE OF INVENTION: KINASE			
FILE REFERENCE: 28113/31941A			
CURRENT APPLICATION NUMBER: US/10/186,399			
CURRENT FILING DATE: 2002-07-01			
PRIOR APPLICATION NUMBER: US 08/320,432			
PRIOR FILING DATE: 1994-10-07			
NUMBER OF SEQ ID NOS: 5			
SOFTWARE: Patentln Ver. 2.0			
SEQ ID NO 2			
LENGTH: 651			
TYPE: PRT			
ORGANISM: Mus musculus			
US-10-186-399-2			
Alignment Scores:			
Pred. No.:	0.0896	Length:	651
Score:	128.00	Matches:	61
Percent Similarity:	39.92%	Conservative:	40
Best Local Similarity:	24.11%	Mismatches:	90
Query Match:	4.19%	Indels:	62
DB:	9	Gaps:	11
US-09-856-061-1 (1-1721) x US-10-186-399-2 (1-651)			
QY	750	CGTGATACACTCCGCCAGAGATGCACACCCCTTACCCGCCAGCACGAGAGAGC	809
Db	141	ProGlyCysThrLeuPrgLValaTYrAlaAspLeuHisIleAlaIleSerAspGluLys	160
QY	810	AGTGCACTTCCTGCCAAGCCACCCTTCCAGAA--GTCCAGAGGGGGCCAGC--	863
Db	161	HisArg-----AlaProThrPheProGluArgLeuLysIleProArgAla	176
QY	864	-----CAGAGAGTGCACAAAAGACTTCAGTAGGCTCCT-----	896
Db	177	ValProValLeuLysMetAspAlaSerSerSerGlyAlaIleLeuPrgGlnTYrAspSer	196
QY	897	-----GGAGCGAAGAGAAATCTCACACACAGACAAAGCCAGAAATCT	938
Db	197	TYrSerLysLysSerCysGlySerGlnProThrSerAsnIleArgTYrIleProArgGlu	216
QY	939	TCCTGCCCATCATCAACCA--AACACACAGAGATGCCACTGCCATTGCCAGCTCT	995
Db	217	AspCysProAspTYrPrgLValaTYrGluLysLysSerGluAspIleAlaI-----	234
QY	996	TCCTACATGCCAGGAAGACACAGATATCAAGCCAGAGACCATACAGTAGCATGCACAC	1055
Db	234	-----	234
QY	1056	TCCTCTGCTCAGAGATGCCAGCTGCAGCCAGCCAGCCCTCGAATGCTGCCCTAT--	1112
Db	235	CysSerAsnGlnLeuGluArgAsnIleAlaSerHisSerThrSerLysMetSerTrpGlu	254

```
OY 1113 -----GAAACACAACCTCGAGAAACCTGACCCACAAAGCCCTGATGAGAGGATGTC 1166
    |||:|||||
Db 255 PheProGluSerSerSerSerGlu-----GluGluGluAsnLeu 267
OY 1167 TGCAGAAATGAAATGCTACATTGGAGAAATACAGTCGCCAGGCAGTGGAGATGTGTTAATG 1226
    |||:|||||
Db 268 HisAlaTyrAspTyrPheAlaGlyAsnIleSerArgSerGlnSerGluGlnLeuLeuArg 287
OY 1227 AAAGAGACAAAGATGTACTTTTGGTCCGAGACTGCTCTACAAATCCAGGCAGAA 1286
    |||:|||||
Db 288 GlnLysGlyLysGluGlyAlaPheMetValArgAsn-----SerSerGlnMetGly 304
OY 1287 CCATATGTTTGGTGGTGT-----TATGGAAACAAGGTCTCAATGTGAAATATC 1337
    |||:|||||
Db 305 MetTyrThrValSerLeuPheSerLysAlaValAsnAspLysGlyThrValLysHis 324
OY 1338 CGTTTCTCGAGAGCAATCACAAGTTGCCCTGGGCACAGGACTTACGAGAAATGAGATG 1397
    |||:|||||
Db 325 TyrHisValHisThrAsnAlaGluAsnLysLeu-----TyrLeuAlaGluAsnTyrCys 342
OY 1398 TTGATTTCTGTGGAAGACATCATGAACTACACATAT 1436
    |||:|||||
Db 343 PheAspSerIleProLysLeuIleHisTyrHisGlnHis 355
```

Search completed: April 21, 2003, 12:35:25
Job time : 58.7798 secs

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:56 ; Search time 53.7435 Seconds
(without alignments)
6156.924 Million cell updates/sec

Title: US-09-856-061-1

Perfect score: 3057
Sequence: 1 acgagcgcccaactgcgcagc.....taaaaaaaaaaaaaaaaaa 1721

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09856061/rnaut_21042003.113019.3758/app.query.fasta_1.3150
-DB=PIR.73 -OEFM=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061.ecgn_1.1.151.ernaut_21042003.113019.3758 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TTIMEOUT=120
-MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: PIR.73.*
2: PIR1.*
3: PIR2.*
4: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	412.5	13.5	533	2	B56110	tyrosine phosphopr
2	398	13.0	533	2	A56110	tyrosine phosphopr
3	158.5	4.9	663	1	TVWVR	protein-tyrosine k
4	148.5	4.9	572	1	T12496	hypothetical prote
5	148.5	4.9	1029	2	T30351	mucin-like protein
6	146	4.8	728	2	H59435	phosphatidylinos
7	145	4.7	723	2	B38749	3-phosphatidylinos
8	140	4.6	1733	2	S27939	tensin - chicken
9	140	4.6	1744	2	A54970	tensin, cardiac mu
10	140	4.6	1792	2	A57075	tensin - chicken (
11	139.5	4.5	1603	2	S17983	gene posterior sex
12	138.5	4.5	443	2	T27877	hypothetical prote
13	138	4.5	1385	2	T21706	hypothetical prote
14	137.5	4.5	592	1	LLBY	actin-binding prot

15	137	4.5	724	2	A38749	3-phosphatidylinos
16	135	4.4	3421	1	WZBER6	367k tegument prot
17	134.5	4.4	480	2	JC7552	Shb-like adapter p
18	134.5	4.4	756	2	JC5886	signaling mediator
19	134	4.4	1599	2	T15854	hypothetical prote
20	133	4.4	675	2	S60612	protein-tyrosine k
21	133	4.4	724	2	A38748	3-phosphatidylinos
22	133	4.4	724	2	A38747	phosphatidylinosit
23	132.5	4.3	559	2	I49444	SH3 binding protei
24	132.5	4.3	649	2	T46500	hypothetical prote
25	132	4.3	673	2	S25776	transforming prote
26	130	4.3	670	2	JC5887	signaling mediator
27	129	4.2	751	2	S68957	adhesin plaque pr
28	128	4.2	706	2	A45990	junctional sarcopl
29	128	4.2	1317	2	T03748	apoptosis associat
30	126.5	4.1	1271	1	TVHUR	ber (breakpoint cl
31	126	4.1	2440	2	S39162	transcription coac
32	125.5	4.1	596	2	I38228	Shb protein - huma
33	125.5	4.1	2437	2	S53611	MIB1 protein - ra
34	125.5	4.1	2774	2	A43359	microtubule-assoc
35	125	4.1	461	2	A57463	p55PIK - mouse
36	125	4.1	856	2	T16543	hypothetical prote
37	125	4.1	962	2	S58107	hypothetical WD-40
38	125	4.1	1210	2	I39410	AF-4 protein, spli
39	124.5	4.1	585	2	A46209	protein-tyrosine-p
40	124.5	4.1	593	1	JN0805	protein-tyrosine-p
41	124.5	4.1	597	1	A53593	protein-tyrosine-p
42	124.5	4.1	879	2	S64755	hypothetical prote
43	123.5	4.0	733	2	S23468	oocyte-specific pr
44	123.5	4.0	841	2	A43254	protein-tyrosine-p
45	123.5	4.0	1736	2	A47747	tight junction pro

ALIGNMENTS

RESULT 1
B56110
tyrosine phosphoprtein SLP-76 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C:Accession: B56110
R:Jackman, J.K.; Molto, D.G.; Sun, Q.; Tanemoto, M.; Turk, C.W.; Peltz, G.A.; Koretz
J. Biol. Chem. 270, 7029-7032, 1995
A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated wit
A:Reference number: A56110; MUID:95221345; PMID:7706237
A:Accession: B56110
A:Status: preliminary
A:Residues: 1-533 <CDS>
A:Residues: 1-533 <ORF>
A:Cross-references: GB:U20159; NID:9806767; PIDN:AGC52189.1; PID:9806768
C:Superfamily: SH2 homology
C:Keywords: phosphoprotein
F:422-520/Domain: SH2 homology <SH2>

Alignment Scores:

Pred. No.: 3,06e-23
Score: 412.50
Percent Similarity: 39.88
Best Local Similarity: 29.66
Query Match: 13.498
DB: 2
Length: 533
Matches: 148
Conservative: 51
Mismatch: 175
Indels: 125
Gaps: 20

US-09-856-061-1 (1-1721) x B56110 (1-533)
QY 294 TTGGGATCTGATTCAGACGCTCTCTCTG-----AAAAATAG 338
||| ||||| |||
Db 57 PheProlyLeuArgMetProLeuLeuSerLysLeuSerLinspLysAnuSanglu 76
QY 339 TCATGGCCAAAGCTCAGACCTGCAAGAGGCGGTGACGGCTTCGACACTTCGCG 398
||||| ||||| |||||
Db 77 GlunArgSerIlePheThrArgLysProGlnIleSerArgPheLeuGluGurGlu 96
QY 399 GATCACAAGAGAACTTGCGTGGG----- 422

```

Db      97  SerHisGluGluAspArgLysIleTrpSerSerPheGluAspAspTyrGluSerProAsn 116
      423  -----GTCCAGGTGGAGAAAATGCAACAGTACACAGCTACGAGATCT----- 470
Db      117  AspAspAspProAspLysIle-----AspAspGlyAspTyrGluSerProAsnGlu 133
      471  GAGTTCAGCTGCTG-----AAGCATGGCCATCA----- 500
Db      134  GluGluGlnAlaLeuValAspAspAlaAspTyrGluIleProProSerAsnAsnGlu 153
      501  -----ATGAAATTTTACCAGCCAGACCTATCCAG-----GATCGGAATAC 542
Db      154  GluAlaLeuGlnSerSerIleLeuProProAsnSerPheHisAsnThrAsnSerMetTyr 173
      543  CGAGATACACGCTATTTCAGATATGAGAGGCTCCCTCTGTACTCT-----CCC 596
Db      174  IleAspArgProProThrGlyLysValSerGlnProProValProProLeuArgPro 193
      597  AAGGCTTCTGTC----- 608
Db      194  LysProAlaLeuProProLeuProThrGlyArgAsnHisSerProLeuSerProIleHis 213
      608  ----- 608
Db      214  ProAsnHisGluGluProSerArgSerGlyAsnAsnLysThrAlaLeuProAlaPro 233
      609  TCCACTGAGAGACAAACAGGAGTGTAGAGATGACACAGCTCGAAGAGATGGAC----- 662
Db      234  SerIleAspArgSerThrLysProProLeuAspArgSerLeuAlaProLeuAspArgLys 253
      663  -----AAGCCTACCTTCAGAGATGTCCAGAACGCCATTAAAGCA 704
Db      254  ProPheIleLeuGlyLysLysProProPheSerAspLysProSerAlaProLeuGlyArg 273
      705  TTCAATATCACAAATAACAGACCTCTTGGCACTCTCGCCCTGCTATCATTCTC 764
Db      274  GlnHisLeuProLysIleGlnLysProProLeuPro-----ProAlaMetAspArg 290
      765  CCCAAGAGTACCAACCTTACCCCCAGCACACAGAGAGAGAGACAGTACTCTGCT 824
Db      291  HisGluArgAsnGluArgLysGlyProValThrThrArg----- 303
      825  CCAAGCCCGCCCTTCCAGAGTCCAGAGGGGGCCCGAGAGAGATGCAAAAGACTTC 884
Db      304  ---LysProSerValProArgHisGlyArgGlyProAspArgArgGlu----- 318
      885  AGTAGGGTCTTGGAGCAGAGAAGATCTCACACACAGCAAGCCAGATCTTCTTGC 944
Db      319  -----AsnAspGluAspAspValHisGlnArgProLeuProGlnProSerLeu 334
      945  CCATCATCAACCAAAACACACAGAGATGTCACCTGCCATTTGCCACTCT----- 995
Db      335  ProSerMetSerSerAsnThrPheProSerArgSerValGlnProSerSerLysAsnThr 354
      996  -----TCTTACATGCCGCGAAGACACAGTATACAGCCAGAGACATACAGGTAGC 1046
Db      355  PheProLeuAlaHisMetProGlyAlaPheSer-----GluSerAsnIleGlyPhe 371
      1047  ATGCAGCACTGT-----CCTGCTCAGAGATGCCAAGCTGCAGCCACAGAGCCCT--- 1097
Db      372  GlnGlnSerAlaSerLeuProProTyrPheSerGlnGlyProGlyAsnAlaArgProLeu 391
      1098  -----CGAATGCTGCCCTATGAAACCAAACTCGAGAAACCTGACCCACCA 1145
Db      392  ArgSerGlnLysArgAsnLeuProLeuProValProAsn-----ArgProGlnProPro 409
      1146  AAGCCTGATGAGAGAGTGC-----TGCGCAATATGAATGTGATGTGAGATACACTGCG 1202
Db      410  SerProGlyLysGluGluThrProLeuAspGluGlnIleTyrValSerTyrIleThrArg 429
      1203  CAGGCACTGGAGATGTGTTAATGAAAGAACAGAGATGTTCTTTTGGTCGAGAC 1262

```

```

Db      430  ProGlnAlaGluAlaLeuAlaLeuArgLysIleAsnGlnIleAspGlyThrPheLeuValArgAsp 449
      1263  TGCTTACAAAATCCAGCAGACCATATGTTTGGGGGTGTTTATGGGAAAGAGTC 1322
Db      450  SerSerLysLysThrAlaAsnAsnProTyrValLeuMetValLeuTyrLysAspLysVal 469
      1323  TACAATGTGAANAATCCGTTTCCCGAGAGCAATCAAGTGTGGCTGGCAGCACTA 1382
Db      470  TyrAsnIleGlnIleArgTyrGlnGlnLysGlnValTyrLeuLeuGlyThrGlyLeu 489
      1383  CGAGAAATGAGATGTTGATTTGTTGGAAGACATCATTAACATCACTATTTTCCC 1442
Db      490  ArgGlyLysGluAspPheLeuSerValSerAspIleIleAspTyrPheArgLysMetPro 509
      1443  ATTCTGTAATAGATGGGAAAGACAGGCTGCACGACGAAAGATGCTACTACCC 1499
Db      510  LeuLeuLeuIleAspGlyLysAsnArgLysSer---ArgTyrGlnCysThrLeuThr 527

RESULT 2
A56110
tyrosine phosphoprotein SLP-76 - human
C:Species: Homo sapiens (man)
C:Date: 28-Apr-1995 #sequence, revision 28-Apr-1995 #text_change 05-Nov-1999
R:Jackman, J.K.; Motto, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretz
J. Biol. Chem. 270, 7029-7032, 1995
C:Accession: A56110
A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated wit
A:Reference number: A56110; MUID:95221345; PMID:7706237
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <JNC>
A:Cross-references: CB:020158; NID:9806765; PIDN:AAC50135.1; PID:9806766
C:Genetics:
A:Gene: GDB: LCP2; SLP-76
A:Cross-references: GDB:1230199; OMIM:601603
A:Map position: 5q33.1-5pter
C:Superfamily: SH2 homology
C:Keywords: phosphoprotein
F:422-520/Domain: SH2 homology <SH2>

Alignment Scores:
Pred. No.: 3, 85e-22 Length: 533
Score: 398.00 Matches: 124
Percent Similarity: 42.93% Conservative: 52
Best Local Similarity: 30.24% Mismatches: 146
Query Match: 13.02% Indels: 88
DB: 2 Gaps: 14

US-09-856-061-1 (1-1721) x A56110 (1-533)
      456  GACTACGAAATCTCTGATCTCCAGCTGCTGAAGCATGGCCATCAATGAAATTTTACCA 515
      144  AspTyrGlnProProProSerAsnAspGluAlaLeuGlnAsn---SerIleLeuPro 162
      516  GCCAGACCTATCCAGAAATCCGAA-----TACGAGATACAGCGATTTTCCAGATATG 569
      163  AlaLysProPheProAsnSerAsnSerMetTyrIleAspArgProProSerGlyLysThr 182
      570  ATGCAGCTGCC-----CTTCTGTATCTCCAGAGCT 602
      183  ProGlnGlnProProValProProGlnArgProMetAlaAlaLeuProProProAla 202
      602  ----- 602
      203  GlyArgAsnHisSerProLeuProProProGlnIleThrAsnHisGluGluProSerArgSer 222
      603  -----TCTGTCTCCACTGAGAGCAAAACGAGGTATG 635
      223  ArgAsnHisLysThrAlaLysLeuProAlaProSerIleAspArgSerThrLysProPro 242
      636  AGAGTACACACCTCGAAGAGTGCAC-----AAGCCTAAC 671

```

Db	243	LeuaspArgSerLeuAlaProheaspArgGluProPheThrLeuGlyLysProPro	262
Qy	672	TTCAAGAT-----GTCAAGCCAAAGCTTTAAAGCATTCATAATACAAAAATA	722
Db	263	PheSerAspLysProSerIleProIaGlyArgSerLeuGlyLuhStLeuProLysIle	282
Qy	723	AACAAGACTCCTTGGCACCCTCTGGCCGTCATACCTCTCCCAAGAATACCAACC	782
Db	283	GlnLysProProLeuProPro-----ThirThruGluHigHisGluArg	296
Qy	783	TTACCCCGCAGCACACAGAGAGAGAGCATGATCTTGCCTCCAAAGCCACTTTTCA	842
Db	297	SerSerProLeuPro-----GlyLysLysProProValPro	308
Qy	843	GAAGTCACAGAGGGGGCCAGCAGAGAGTGCAAAAGACTTCATAGGTCTTGAGACA	902
Db	309	LysHisGlyTrpGlyProAspArgGlu-----AsnAsp	320
Qy	903	GAAGAAGAATCTCACACACAGACAAAGCCAGAA-----TCTTCTTGC	944
Db	321	GluAspAspValHisGlnArgProLeuProGluProIaLeuProMetSerSerAsn	340
Qy	945	CCATCATCAACCAACAAACACACAGAGAGTCCACCTGCCATGTCCAGCTTTCCTCATG	1004
Db	341	ThrPheProSerArgSerThrLysProSerProMetAsnProLeuProSerSerHisMet	360
Qy	1005	CCAGAGAACGACAGTATACAAAGCCAGAGACCATACAGGTATGACATGACACACTGTCTCT	1064
Db	361	ProGlyAlaPheSer---GluSerAsnSerSerPheProGluSerAlaSerLeuProPro	379
Qy	1065	CAGAGATGCCACAGCTGCAGCCAGCCACAGCCT-----CGAATGCTGCC	1109
Db	380	TyrPheSerGlnGlyProSerAsnArgProProIleArgAlaGluGlyLysAsnPhePro	399
Qy	1110	TATGAACAACAACCTGGAGAAACCTGACCCCAAAAGCCTGATGAGAAGATGTCTGG	1169
Db	400	LeuProLeuProAsnLysProArgPro---ProSerProIaGluGluGlnAsnSerLeu	418
Qy	1170	CAGAAATGAATGTACATGTGAGAAATACAGTCCAGCAGTGAGAGATGTTAATGAAA	1229
Db	419	AsnGluGluTrpLysValSerTyrIleThrArgProGluAlaGlnAlaLeuArgLys	438
Qy	1230	GAGAACAGAGATGGTACTTTTGTGTCGAGACCTCTACAAAATCCAAAGCAGACAACA	1289
Db	439	IleAsnGlnAspGlyThrPheLeuValArgAspSerSerLysThrThrAsnPro	458
Qy	1290	TATGTTTGGTGTGTTTATGTGGAACAAGGTCTACAATGTGAAAATCCGTTTCTCGAG	1349
Db	459	TyrValLeuMetValLeuTyrLysAspLysValTyrAsnIleGlnIleArgTyrGlnLys	478
Qy	1350	AGCAATCAACAGTTTGCCTTGGCAGACAGACTAGCAGGAATAGAGATGTTTGTGTG	1409
Db	479	GluSerIleValTyrLeuLeuGlyThrGlyLeuArgGlyLysGluAspPheLeuSerVal	498
Qy	1410	GAAGACATCATTTGAGACATACACATATTTCCATTTGCTAATAGATGGAAAGACAAG	1469
Db	499	SerAspIleIleAspTyrPheArgLysMetProLeuLeuLeuIleAspLysAsnArg	518
Qy	1470	GCTGCACGCGAAGAAACAGTGTACTCAC	1499
Db	519	GlySer---ArgTyrGlnCysThrLeuThr	527

RESULT 3
TVWMR
protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rashevsky)
C:Species: feline sarcoma virus
A:Note: host Felis sp. (cat)
C:Date: 27-Nov-1995 #sequence_revision 26-May-1995 #text_change 31-Mar-2000
C:Accession: A00653; A03937
R:Naharro, G.; Robbins, K.C.; Reddy, E.P.
Science 223, 63-66, 1984
A:Title: Gene product of v-fgr onc: hybrid protein containing a portion of actin and a tyrosine kinase domain
#:Reference number: A00653; M0UD:84097512; PMID:6318314

A:Accession: A00653
A:Molecule type: DNA
A:Residues: 1-663 <NAH>
A:Cross-references: GB:X00255; GB:K01487; NID:961542; PID:961543
A:Note: the authors translated the codon GAR for residue 14 as Glu
C:Comment: This protein is synthesized as a gag-fgr polypotein.
C:Genetics:
A:Gene: fgr
C:Superfamily: feline sarcoma virus protein-tyrosine kinase fgr; protein kinase homol
C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; p
F:1-118/Region: gag polypotein similarity
F:141-268/Region: actin similarity
F:285-382/Domain: SH2 homology <SH2>
F:402-660/Domain: protein kinase homology <KIN>
F:410-418/Region: protein kinase ATP-binding motif
F:432/Active site: Lys #status predicted

Alignment Scores:

Pred. No.:	0.000589	Length:	663
Score:	158..50	Matches:	74
Percent Similarity:	35.57%	Conservative:	32
Best Local Similarity:	24.83%	Mismatches:	96
Query Match:	5.18%	Indels:	97
DB:	1	Gaps:	13

US-09-856-061-1 (1-1721) x TVMVR (1-663)

QY	735	TTGCACCTCTCGGCGCTGCTATCACTCTCCCAAGAGTAC	776
Db	101	LeuProProProLysProProThrSerLeuProGlnProHisSerProGlnProAlaArg	120
QY	777	-----CAACCTTACCCCGACACACCA	800
Db	121	AlaLeuysArgrProAlaValCysArgrProLeuProIleuProThrAla	140
QY	801	-----GAGAGAGCACTGCATCTTC	821
Db	141	MetGlnGlnGlnValAlaAlaLeuValIleAspAsnGlySerGlyMetCysValaGly	160
QY	822	-----GTCCAAAGCCACCTTCCAGAAAGTCCAGAGGGGCCA	861
Db	161	PheAlaGlyAspAspAlaProArgrAlaValAlaPheProSerIleVal-GlyArgrProArgrH	180
QY	862	-----CGCAGAGACTGCAGAAAGACTTCAGTAGAGGCTCTTGAGACGAA	905
Db	180	SerGlnGlyValMetValCysGlyMetGlyGlnLysAspSerTyrValGlyAspGlnAlaGlnSe	200
QY	906	GAGAAAT-----CTCACCCACGACA	926
Db	200	ArgArgrGlyIleLeuThrIleuLysTyrProIleGlnHisGlyIleValAlaThrAsnTrpAs	220
QY	927	AAGCCAGATCTTCTCCCATCATCAACCAAAACACACACAGACAGTCCACCTGCCATT	986
Db	220	PAspMetGlnLysIleThrPheHisThrPheTyrAsnGlnLeuAlaGVal-----	236
QY	987	GCCAGCTCTCTCATGCGCAGGAAAGACACAGTATCAAGCCACGAGACCATACAGGTAGC	1046
Db	237	-----AlaProGlnGlnL-----	240
QY	1047	ATGCAGCACTCTCTGCTCAGAGATGCCAAGCTGCAGCCACGACACAGCCCTCGAATG--	1103
Db	241	-----HisProValIleuLeuThrGlnAlaProLeuAsnProLysAlaAsnArgGlyLysMe	259
QY	1104	-----CTGCCCATATGAAACACAAACACTGGAGAAACCTGCACCCACCAACCTCGATGA	1156
Db	259	ThrGlnIleMetThrGlnThrPheAsnIle-----ProSerAsnTyrValAlaProVa	277
QY	1157	GAAAGATCTCTGGCAGAAATAGTGTATCTTGAGAGATACAGTCCGACAGCAGTGGAA--	1214
Db	277	LAspSerIleGlnAlaGlnGlnTyrPyrIleGlyLysIleGlyValArgLysAspAlaGlnTr	297
QY	1215	-CATGTGTTAATGAAGAAGACAGAGATGTACTTTTGGTCCGAGACTGCTTCACAAA	1273

Db	297	gslnleuLeuSerProGlyAseAlaArgGlyAlaIlePheLeuValAlaGluSerCiluthrTh	317
QY	1274	ATCCAAAGCAGACACCATATGTTTGGTGGTG-----TTTTATGGGAAACA	1318
Db	317	rlYsglYAla-----TyrSerLeuSerTleIArgAspTTPdSPdLAlaArgGlyAspI	335
QY	1319	GGTCGTACATGTGAATAATCCGTTTCCTCGAGAGCAATCAACACTTTGCCCTGGGCACAG	1378
Db	335	svAlaYhNstYrLysIleAlaGlyLsleAspThGly-----GlyTYrTY	350
QY	1379	ACTACGAGCAATATGAGATGTTTGAATTCTGTGGAGAGACATCATTAACACTAC	1430
Db	350	rlleThrThrArgAlaIlePheAsnSerValIGlnGluLeuValGlnHstYr	367
RESULT 4			
T12496			
hypotheetical protein DKFZp434I091.1 - human (fragment)			
C:Species: Homo sapiens (man)			
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000			
C:Accession: T12496; T46492			
R:Poultka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.			
submitted to the Protein Sequence Database, June 1999			
A:Reference number: 217525			
A:Accession: T12496			
A:Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-572 <POU>			
A:Cross-References: EMBL:AL080174			
A:Experimental source: adult testis; clone DKFZp434I091			
R:Poultka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.			
submitted to the Protein Sequence Database, January 2000			
A:Reference number: 223029			
A:Accession: T46492			
A:Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 456-572 <AAA>			
A:Cross-References: EMBL:AL137328			
A:Experimental source: adult testis; clone DKFZp434A2372			
C:Genetics:			
A:Note: DKFZp434I091.1; DKFZp434A2372.1			
Alignment Scores:			
Pred. No.:	0.00276	Length:	572
Score:	149.50	Matches:	106
Percent Similarity:	33.59%	Conservative:	66
Best Local Similarity:	20.70%	Mismatches:	200
Query Match:	4.89%	Indels:	140
Db:	2	Gaps:	25
US-09-856-061-1 (1-1721) x T12496 (1-572)			
QY	240	CAGGTGCAGACGACCATGACACGCCAGGCCATTAAGGACACAGAA--GAAGATTC	296
Db	131	GlnleuProProGlnThrArgProAlaGlyAsnLysAspGlnGlnGlyTyrGlnGlyPhe	150
QY	297	GGTGATCTGAGATTCGAGAACGTCGTCGCTGAGAAATAAGTCATGAGCCAGCCTCAGC	356
Db	151	-----serIleIleSerAnGlnPoleuAlaIleThrYrGln	162
QY	357	AGTGCACAAAGGGGGGTGTGAGCGGTGTGGAACACACTTCGGATTCACAGAGAACTTG	416
Db	163	GluAsnArgAsnThrYsPheSer-----ProAspHstLysGlnProGln	177
QY	417	GCTGGGGTCCACAGTGTGAGAAAAA-----TTCACACACTTAACACAGACTCAAGAAATCC	470
Db	178	ProGlyThrAlaGlyLysMetArgValSerValSerSerLysGlySerTleSerThrPro	197
QY	471	GAGTTCACAGTGTGCTGAAGGCAATGCGCATCAATGAATAATTTCACAGCC-----AGACCT	524
Db	198	AsnSerProVal-----LysMetThrLeuIleProSerProAsnSerPro	212
QY	525	ATCCAG-----GAATGGAAATACGACGATACA	551

Db	213	PhgEluYsValGlYsLeuAlaSerSerSphTrGlYsLeuSerApgInSerSerThr	232
QY	552	CGGTATTTCCAGATATATATGAGAGCTCCCTTCTGTTAACTCCCAAGCTTCTGTCTC	6111
Db	233	-----GluThrAspSerThrTrpYsSerGlnGluSerAsn	245
QY	612	ACTGAGAGACAAACCGAGGATGTGAGATGACACAGCTGGAGAAAGTGGACAAAGCCATAC	6711
Db	246	ProLYsLeuAspProGlnLeuLeuAlaGlnYsIleLeuGlnGluTrpGlnSerHisLeu	265
QY	672	TTTCAGGATGTCAAGACCAACGCTTTAAAGATTCAATACAAATAATTAACAAAGCT	731
Db	266	IleAlaValaGluArgLeuGlnArgSerGlyGly	281
QY	732	CCATTGCCACT	743
Db	282	AsnAspProGluAspGlyValaGlnAlaProSerSerThrAlaValaPheArgAlaSerGlu	301
QY	744	-----CCCTGGCTGTCTATCACTCTCCCAAGAAAGTACCAACCTTACCC	788
Db	302	ThrSerAlaPheSerArgProValLeuSer	320
QY	789	-----CCAGCACACACAGAGACGACGTCCATCTTCCTCCAAAGCCACCTTT	839
Db	321	ThrValYsProLYsProLYsProAlaArgSerSer	333
QY	840	CCAGAAAGTCCAGAGGGGGCCAGGACAGAGGATGCCAAAGACTTCAGTAGGCTTGGCA	899
Db	334	ProLYsValSerSerGly	350
QY	900	GCAGAAAGAAATCTCAACCAACAGACAAAGCCAGAAATCTTTGCCATCATCAACCA	959
Db	351	LYsAspSerProSerGlnHisSerGlyArgProSerProGlyYsAspSer	368
QY	960	AACACACAGAGAGAGCCACCTGCCATTGGCCAGCTCTCCACATGGCCAGAAACAGCAT	1011
Db	369	ThrSerGlnLeuAspGlnPro	387
QY	1020	ATACAAAGCCAGACACCATACAGTACATGACATGCTCTGCTCCAGATGGCCAACT	1071
Db	388	AlaHisIleSerLYsSerProArgAsnMetSerProSerSerGlyHisGlnSerProAla	407
QY	1080	GCAGGCACCCACACCCCTCGAATGCTCTCCCTATGAAACACAACTCGGAAACCTGAC	1133
Db	408	GlySerAlaProSerPro	426
QY	1140	CCACACA--AAGCTGATGGAAGAGATGTCTGGCAATGAATGATCAATTGACAAATAC	1196
Db	427	ProAlaAspAlaProAsp	436
QY	1197	AGTGGCCAGGACAGTGAAGATGTGTTAATGAAGACAAACAGAGATGCTATTTGGTCT	1256
Db	437	LYsMetAlaAlaIleAspGluLYsValaGlnAlaValaHis	456
QY	1257	CGAAGTGTCTTACAAATATCAAGGACGACCAATATGTTGTGGTGATGTTATGGAC	1316
Db	456	AspThr	472
QY	1317	AAGGTCTACATGTGAATAATCCCTTTCTCGACAGCAATCAACAGTTTG	1365
Db	472	AspGlyThrMet	489
QY	1366	-----CCCTGGGCACAGACGTACGAGAAATGACATGTGTTGATTTCT	1406
Db	489	TrpArgHisAsnLYsLeuGluGluLYsValaLYsProLYsLeuGluLeuLYsGluLeuSerLe	509
QY	1407	G--	1421
Db	509	UGAGlnHisAspGlyAlaArgProGluAlaArgProAsnGlnYsIleTrpArgTrpGlnGlu	529
QY	1422	GAAACATACACATATTTCCCATCTTGTGTAATAGATGGGAAAGACAGCTGACAGG	1481
Db	529	TrpHisSer	540

QY 1482 AACACTGCTACCTACCCAGCCATGCTC-----TGCCAGGCTCTTCACT 1532
 Db 540 yProProAlaThrAlaProAlaArgProLeuArgProSerGlyAsnGlyTyrIysPh 560
 QY 1533 CAGTACTCCAGCCAGCATTTCATGAGTAAGAAG 1566
 Db 560 eLeuSerProGlyArgPhePheProSerSerIys 571

RESULT 5

T30351
 mucin-like protein - lymphatic virus
 C:Species: Lymphatic virus
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T30351
 R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohu
 Virology 253, 17-34, 1999
 A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for lymphatic
 A:Reference number: Z20836; PMID:99124785; PMID:9887315
 A:Accession: T30351
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1029 <K02>
 A:Cross-references: EMBL:AF081810; PIDN:AC70189.1

Alignment Scores:

Pred. No.:	0.00365	Length:	1029
Score:	148.50	Matches:	101
Percent Similarity:	34.99%	Conservative:	75
Best Local Similarity:	20.08%	Mismatches:	217
Query Match:	4.86%	Indels:	110
DB:	2	Gaps:	18

US-09-856-061-1 (1-1721) x T30351 (1-1029)

QY 321 TCCTGCTGAAATAGTATGCGCAAGCTTCAGAGTGCAGGCGGTGCGAGCG 380
 Db 495 SerThrValysProLysSerSerProAspLysSerProAspLysSerSerSer 514
 QY 381 GTTCTGAGACCACTTCGCGATCAGAGAGCACTTGGCTGGGCTCCAGGTGGAGAAA 440
 Db 515 ThrValysPro-----LysSerSerSerAlaValProAspLysProLys 529
 QY 441 TGCACAGTACAGCAGTACAGAA-----GATCCTGAGTTCAGCTGAGAGCA 491
 Db 530 HisSerSerAspLysSerAlaValThrValysProLys----- 543
 QY 492 TGGCATCATGAATTTTACAGCAGCAGCTATCCAGAAATGGAATACGAGATACA 551
 Db 544 -----SerSerSerAlaValProAspLysProSerAspLysProSerSerAspLys 561
 QY 552 CGCTATTTCAGGATGATGAG----- 575
 Db 562 ArgSerPheArgProProSerGlySerGlyArgAspGlnProLysSerLeuAlaGlnThr 581
 QY 576 GCTCCCTTCTGTTACCTCCAGAGCTTCTCTCAGTACAGAGCAAAAGGAGATGTG 635
 Db 582 SerProPheLeuAlaAlaAlaLeuAlaArgAspMetAspGlnArgLeuAsnArgAsp 600
 QY 636 AGATGACACAGCTGGAAGAAGTGAGCAAGCTTCAAGATGTCAGAGCAAGCCG 695
 Db 601 -----LeuGlnAlaAlaValAlaSerIleThrGlyGlyValArgProGlnPhe 616
 QY 636 TTTAAAGG-----TTCAATACACAAAATTAACAAGCT----- 731
 Db 617 TyrProAlaProGlnProSerAlaProValAspTyrPheThrLeuAsnSerAlaGlnPhe 636
 QY 732 -----CCTTTCAGACCTCTCGCGCTGATACCTCCCAAGAAAGTAC----- 776
 Db 637 AlaProGlnProAlaProGlnProValProAlaProGlnProSerAlaProValAspTyrPheThr 656
 QY 777 -----CAACCTTACCC-----CCA 791

Db 657 LeuAsnSerAlaGlnPheAlaProGlnProAlaProGlnProAlaProGlnProAlaPro 676
 QY 792 GCACACACAG 851
 Db 677 GluProAlaProGlnProThrSerArgPheThrSerGlnProAlaLeuGlyProValGlu 696
 QY 852 AGGGGG-----CCAGGAG 905
 Db 697 ProAlaLeuGlnProValGlnProAlaLeuGlnProValGlnProAlaLeuGlnProGly 716
 QY 906 GAGATCTCACCACAGCAAGCA-----GAATCTTCTGCCATCA 950
 Db 717 GluProAlaLeuGlnProGlyGlnProAlaLeuGlnProGlyGlnProAlaLeuAla 736
 QY 951 TCACACCAAAACACAG 998
 Db 737 AlaSerGlnLeuAlaValGlnArgProAlaGlnProAlaProAspLeuThrSerSer 756
 QY 999 TACATCCAGAAAGCAGATATACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1058
 Db 757 AlaVal-----GluSerThrPheGlyHisLysSerThrValAlaSerGlnLeuAla 773
 QY 1059 CCGTCTCAGAGATGCCAGAGCTGACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1118
 Db 774 AlaArgProAlaAlaGlnLeuAlaAlaAlaProAlaProGlnLeuThrProGlnHisSer 793
 QY 1119 ACAACCTCGAGAGAA-----CCT 1136
 Db 794 GluProAlaProGlnProAlaSerLysProAlaSerGlnProSerAlaHisLeuThrPro 813
 QY 1137 GACCCCAAG 1196
 Db 814 AspArgThrLeuProSerProLeuThrValLysSerLeuGlnSerLeuLysSerAsp 833
 QY 1197 AGTCCAG 1253
 Db 834 AspGlnGlnValAlaGlnAspIlePheValLysAsnGlnAspGlnLysProThrValGlu 853
 QY 1254 GTCCGAGACTGCTCTACAAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1313
 Db 854 ValGlnAspIleThrIleLysAsnGlnAsnAsnGlnLysLeuValGlyValProGly 873
 QY 1314 AACAGGCTACAAATGGAATCCGCTTCTCGAGAGCAATCAACAGTTCGCCCTGGGC 1373
 Db 874 -----AspArgLysHisLysAlaSerSer 882
 QY 1374 ACAGAGTACAGAGAAATGAGATGTTGATTCGTGGAACATCAAT-----GAA 1424
 Db 883 ThrProSerAlaSerAlaGlnValAlaAspAspGlnValGlnIleMetSerPheProGlu 902
 QY 1425 CACTACACATATTTCCCATTTCTGCTAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1469
 Db 903 AsnHisLysAlaLysProGlyValPheMetLeuAsnLysArgLysSerAspLeuMet 922
 QY 1470 GCTGACGAGAGAAAGAGTCTCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1529
 Db 923 SerValArgLysArgProArgTyrSerGlnGlnProGlnProProProProPhePhe 942
 QY 1530 ACTGAGTACTCAGCAG 1589
 Db 942 OProGlnPheProProTyrSerPheAlaSerProTyrAlaAlaSerAlaSerAlaHisGln 962
 QY 1590 GTGGCCT 1596
 Db 962 nTyrPro 964

RESULT 6

H59435
 phosphoinositide-3-kinase regulatory beta chain [imported] - human
 C:Species: Homo sapiens (man)
 C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: H59435; A59436
 R:Volinia S; Patrachini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke I


```

OY 506 AATTTTACGACCGACCTATCCAGGAATCGGAATACGACATATTCCTCCAGCA 565
    : ||||| ||||| ||||| :
Db 94 gProleuProAlaArgPro--ArgAspGlyProProGluProGlyLeuThrLeuProAs 113
OY 566 TTGTGATGAGG-----GCTCCCTTCGTGTACCTCC-----AAGGCTTCTGT 607
    : ||||| ||||| ||||| :
Db 113 pLeuProGluGlnPheSerProProAspValAlaProProIleLeuValLysLeuValG1 133
OY 608 CTCCTAGAGAGACAACCGAGGAT-----GTGAGG----- 638
    : ||||| ||||| ||||| :
Db 133 uAlaIleGluArgThrGlyLeuAspSerThrArgProGluProProAlaValArgThrAs 153
OY 639 -----ATGACACAGAGCTGGAGAGAGTGCACAAAGCTTACAGAGAGTGCAGAGCA 691
    : ||||| ||||| ||||| :
Db 153 pTTPSerLeuSerAspValGluGlnIleTrpAspAlaAlaLeuSerAsp----- 169
OY 692 AGCCTTAAAGATCAATACACAAATAAACAAGACTCTTTCGCACTCTCCGCGCC 751
    : ||||| ||||| ||||| :
Db 170 -GlyValLysGlyPheLeu-----LeuAlaLeuProAlaPr 181
OY 752 TCGTATCACTCTCCCAAGAACTA----- 775
    : ||||| ||||| ||||| :
Db 181 oLeuValThr--ProGluAlaAlaAlaGluAlaHisArgAlaLeuArgGluAlaAlaGly 200
OY 776 -----CCAACTTACCCCGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 823
    : ||||| ||||| ||||| :
Db 201 ValGlyProAlaLeuGluProProThrLeuProLeuHisAlaLeuThrLeuAlaGlyPhe 220
OY 824 TCCAAAGCCCACTT-----TCCAGAGTCCAGAGAGGAGGAGGAGGAGGAGAGAGAG 871
    : ||||| ||||| ||||| :
Db 221 LeuLeuGlnHisLeuGlyArgValAlaGlyArgAlaProAlaProGlyProAlaValArg 240
OY 872 TGCAAAAGACTTACAGAGAGGCTCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 931
    : ||||| ||||| ||||| :
Db 241 AlaLeuGlyAlaThrPheGlyProLeuLeuLeuArgAlaProProProPro--SerProPr 260
OY 932 AGAATCTCTTGGCCATCATCAACAAACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 991
    : ||||| ||||| ||||| :
Db 260 oProGlyGlyAlaProAspGlyThrGluProThrProAspPheProAlaLeuLeuValG1 280
OY 992 CTCCTTCCATAGCCAGAGAAAGCAGACTATACAGAGAGAGAGAGAGAGAGAGAGAGAG 1051
    : ||||| ||||| ||||| :
Db 280 uLysLeuLeu----- 1151
OY 1052 GCAGTCTCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1111
    : ||||| ||||| ||||| :
Db 285 uHisLeuGluGluGlnGluVal-----AlaProProAlaLeuPro-- 298
OY 1112 TGAAGAACAAACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1151
    : ||||| ||||| ||||| :
Db 299 -----ProLysProProLysThrLysProAlaProThrGlyLeuAlaAs 313
OY 1152 -----GATGAGAAGAGATGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1204
    : ||||| ||||| ||||| :
Db 313 nGlyGlySerProProSerLeuGlnAspAlaGluTrpTyrTrpLysPheSerArgG1 333
OY 1205 GGCAGTGGAGAGATGTTTATGAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1264
    : ||||| ||||| ||||| :
Db 333 uGluValAsnGluLysLeu--ArgAspThrProAspGlyThrPheLeuValArgAspAl 352
OY 1265 CTCACAGAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1324
    : ||||| ||||| ||||| :
Db 352 aSerSerLysIleGlnGlyLys--TyrThrLeuThrLeuArgLysGlyLysAsn----- 369
OY 1325 CAATGTGAAATCCGTTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384
    : ||||| ||||| ||||| :
Db 370 -AsnLysLeuLysValPheHisArgAspGlyHisIleGlyPheSerGluProLeu-- 388
OY 1385 AGGAATAGAGATGTTTGTGATCTGTGAGAGAGAGATCTTAAACACTACATAT 1436
    : ||||| ||||| ||||| :
Db 389 -----ThrPheCysSerValValAspLeuIleThrHisTyrArgHis 402

```

```

RESULT 8
S27939
tensin - chicken
C:Species: Gallus gallus (chicken)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 12-Feb-1999
C:Accession: S27939; S28973
R:Chen, L.B.
submitted to the EMBL Data Library, August 1991
A:Reference number: S27939
A:Accession: S27939
A:Molecule type: mRNA
A:Residues: 1-1733 <CHE>
A:Cross-references: EMBL:M74165; NID:9212751; PID:9212752
R:Weigt, C.; Gaertner, A.; Wegner, A.; Korte, H.; Meyer, H.E.
J. Mol. Biol. 227, 593-595, 1992
A:Title: Occurrence of an actin-inserting domain in tensin.
A:Reference number: S28973; MUID:93021103; PMID:1404377
A:Accession: S28973
A:Molecule type: Protein
A:Residues: 862-871, 'X', 873-875, 'A', 877-1212 <WEI>
C:Superfamily: SH2 homology
P:1461-1570/Domain: SH2 homology <SH2>

```

```

Alignment Scores:
Pred. No.: 0.0177 Length: 1733
Score: 140.00 Matches: 93
Percent Similarity: 34.11% Conservative: 54
Best Local Similarity: 21.58% Mismatches: 138
Query Match: 4.58% Indels: 146
Gaps: 21
US-09-856-061-1 (1-1721) x S27939 (1-1733)

```

```

OY 527 CCAGGAATGGGAATAGCGATACAGCTATTTCCAGATATGATGAGAGCTCCCTCT 586
    : ||||| ||||| ||||| :
Db 1288 ProGlyLeuGly-----ArgHisThrValSerProHisAlaProProGlySerProSer 1305
OY 587 GTTACCTCCCA----- 610
    : ||||| ||||| ||||| :
Db 1306 LeuAlaArgHisGlnMetAlaAlaValProProGlySerProMetGlyGlyTyrSer--Se 1325
OY 611 CACTGAGAGACAACACAGGAGATGAGAGATGACACAGCTGGAAGAGTGGACAGCTTAC 670
    : ||||| ||||| ||||| :
Db 1325 rProGluGluArg-----ArgProThr 1332
OY 671 CTCACAGGATGTCAGAGAGCCAGCTTAAAGATCAATACAAATAAACAAGAG 730
    : ||||| ||||| ||||| :
Db 1332 rLeuSer-----ArgGlnSerSerAlaSerGlyTyrGln----- 1343
OY 731 TCCTTTGCACCTCTCGGCTCTATACCTTCCCAAGAGATGCCAACCC----- 782
    : ||||| ||||| ||||| :
Db 1344 -----ProProSerThrProSerPheProValSerProAlaTyrTrpProGlyThrSe 1361
OY 783 -TTACCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 829
    : ||||| ||||| ||||| :
Db 1361 rThrProHisSerSerSerProAspSerAlaAlaTyrArgGlnGlySerProThrProG1 1381
OY 830 GCCCAGCTTCCAGAGAGTCCAGAG-----GGGCCAGAGAGAGAGAGAGAGAGAGAGAG 880
    : ||||| ||||| ||||| :
Db 1381 nProAlaLeuProGluLysArgArgGlyMetSerAlaGlyGluArgSerAsnSerLeuProAs 1401
OY 881 CTCCTAGAGGCTCTTGGAGAGAGAGAGAGAGAGATTCACACACAGAGAGAGAGAGAG 940
    : ||||| ||||| ||||| :
Db 1401 nTyrAlaThrValAsnGly-----LysAlaSe 1410
OY 941 TTGCCCCA-----TCATCAACCAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 997
    : ||||| ||||| ||||| :
Db 1410 rSerProLeuSerSerGlyMetSerSerProSerSerGlySerAlaValAlaPheSerH1 1430
OY 998 CTCACAGGCA-----GGAAGACACAGTATACAGAGAGAGAGAGAGAGAGAGAGAGAG 1051
    : ||||| ||||| ||||| :
Db 1430 sThrLeuProAspPheSerLysPheSerMet----- 1440

```

QY 1052 GCACGTCTGCTGCTCAGAGATGCGAAGCTGACGACGACCCCTGCAATGCTGCTCA 1111
 Db 1441 -----ProaspiliserProgl 1446
 QY 1112 TGAACAACAACTCGGAAACCTGACCCACAAAGCTGTAGAGATGCTGCGCA 1171
 Db 1446 uThrArgAlaAsnValLysPheValGlnAspThr-----Se 1458
 QY 1172 GAATGAATGCTACATTTGGAGATACAGTCCG---CAGCGAGTGAAGATGTTAATGAA 1228
 Db 1458 rlystyrtlyrtyrlyspProaspiliserArgLuglnAlaIle-----AlaLeuLeuLys 1476
 QY 1229 AGAGAACAAAGATGCTACTTTTGGTCCGACAGCTGCTTACAAATGCAAGCAGAAC 1288
 Db 1476 sasprArgLuglnProglYAlaPheIleIleArgspSerHisSer----- 1490
 QY 1289 ATATGTTTTGGTGTGTTTATATGCAACAAAGCTCTACAAATGCAATCCGTTTCTCGA 1348
 Db 1491 -----PheArgGly-----AlaTyrGlyLeuAlaMetLysValAlaLase 1503
 QY 1349 GAGCAATCAACAGTTGGCCCTGGCAGACGACTACGAGAAATGAGATGTTGATCTGT 1408
 Db 1503 rProProPro-----ThValMetGlnGlnAsnLysLysGlyAspIleTh 1518
 QY 1409 GGAACACATCATTTGACACTACATATATTTCCATTCCTTAATAGTGGAAAGACAA 1468
 Db 1518 rAsnGlnLeuValArgHisPhe-----LeuIleGlnThrSerProAr 1532
 QY 1469 GGCTCCAGCAGAAACAGTGC----- 1490
 Db 1532 gGlyValLysLeuLysGlyCysProAsnGlnProAsnPheGlyCysLeuSerAlaLeuVa 1552
 QY 1491 -TACCTCACCAGCCACTGCTCTGCGCAGGCTCTTCTCACTAGTACTCCAGCAGGC 1549
 Db 1552 lTyrGlnHisSerIleMetProLeuAlaLeu-ProCysLysLeuValIleProAspArg 1572
 QY 1550 ACTTCAGTAAAGACCCAGCCAGACAGATATGCCCGCATCACTGGCCCTGTCTCA 1609
 Db 1572 sPromeGlnGlnLysLysAspAlaLase----- 1582
 QY 1610 TTCCTGCTCAATGATGATCACTCTTCTTCATCTGATTCGCAATATATATTTT 1669
 Db 1583 --ThrThrAsnSerAlaThrAspLeuLeuLysGlnGlyAlaAlaCysAsnValLeuTher 1602
 QY 1670 CTGTGCTTCAAGGATGATTTTGTACT 1698
 Db 1602 leAsnSerValGlnMetGlnSerLeuThr 1611
 RESULT 9
 A54970
 tensin, cardiac muscle - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 11-Nov-1994 #sequence, revision 11-Nov-1994 #text, change 21-Jul-2000
 C:Accession: A54970; S38330; S21544
 R:Do, S.H.; An, Q.; Bao, S.; Wong, W.K.; Liu, Y.; Janney, P.A.; Hartwig, J.H.; Chen, L.H.
 J. Biol. Chem. 269, 22310-22319, 1994
 A:Title: Molecular cloning of chick cardiac muscle tensin. Full-length cDNA sequence, ex
 A:Accession number: A54970; MIMD:9430987; PMID:8071358
 A:Reference number: A54970
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-1744 <LOA>
 A:Cross-references: GB:M96625
 R:van de Werken, R.; Gennari, M.; Tavelle, S.; Bel, P.; Molina, F.; Liu, S.; Canceda, F
 Eur. J. Biochem. 217, 781-790, 1993
 A:Title: Modulation of tensin and vimentin expression in chick embryo developing cartila
 A:Reference number: S38330; MIMD:94039118; PMID:8223621
 A:Accession: S38330
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1469-1744 <VAP>
 A:Cross-references: EMBL:X66286; MIMD:963802; PIDN:CAA46992.1; PID:963803
 C:Superfamily: SH2 homology

C:Keywords: cardiac muscle; heart
 F:1472-1581/Domain: SH2 homology <SH2>
 Alignment Scores:
 Pred. No.: 0.0177 Length: 1744
 Score: 140.00 Matches: 93
 Percent Similarity: 34.11% Conservative: 54
 Best Local Similarity: 21.58% Mismatches: 138
 Query Match: 4.58% Indels: 146
 DB: 2 Gaps: 21
 US-09-856-061-1 (1-1721) x A54970 (1-1744)
 QY 527 CCAGAAATCGGAATACGAGATACACGCTATTTCCAGGATATGATGAGCGCTCCCTCT 586
 Db 1299 ProGlyLeuGly-----ArgHisThrValSerProHisAlaProProGlySerProSer 1316
 QY 587 GTTACCTCCCA----- 610
 Db 1317 LeuAlaArgHisGlnMetAlaAlaValProProGlySerPromeTyrGlyTyrSer--Se 1336
 QY 611 CACTGAGACAAACACGAGATGTAGAGATACACAGCTGGAAAGTGGCAAGCCTAC 670
 Db 1336 rProGlnGlnArg-----ArgProth 1343
 QY 671 CTTGAGAGATGTCAGAAACCCCTTAAAGCTTCAATTAACAAATAAACAAGAC 730
 Db 1343 rLeuSer-----ArgLysSerSerAlaSerGlyTyrGln----- 1354
 QY 731 TCCCTTCCACCTTCCTGCGCTGTATCACTTCCGCCAAGAACCAAGCAACCC-- 782
 Db 1355 -----ProProSerThrProSerPheProAlaTyrTyrProGlyLysrse 1372
 QY 783 -TTACCCCAAGCAGCAGACAGAGAGACAGTGCATCTTC-----GCTCCAA 829
 Db 1372 rThrProHisSerSerSerProAspSerAlaAlaTyrArgGlnGlySerProThrProgl 1392
 QY 830 GCCCACCCTTTCAGAAATCCAGAG-----GGGCCAGCAGAGAGTGCACAAAG 880
 Db 1392 nProAlaLeuProGlnLysArgArgMetSerAlaGlyGlnArgSerHisSerLeuPro 1412
 QY 881 CTTGAGTACGCTCTGAGCAGACAGAAAGATCTCACACCAACAAAGCAGATCTTC 940
 Db 1412 nTyrAlaThrValAsnGly-----LysLase 1421
 QY 941 TTGCGCA--TCATCAACCAAAACACAGAGATCCACCTGCCATGCCAGCTCTTC 997
 Db 1421 rSerProLeuSerSerGlyMetSerSerProSerSerGlySerAlaValAlaPheSerHis 1441
 QY 998 CTACATGCCA-----GGAACCAACAGTATACAAACGACAGACCAATACAGTATGCA 1051
 Db 1441 sThrLeuProAspPheSerLysPheSerMet----- 1451
 QY 1052 GCACGTCTGCTGCTCAGAGATGCGAAGCTGACCCACAAAGCTGTAGAGATGCTGCGCA 1111
 Db 1452 -----ProAspIleSerProgl 1457
 QY 1112 TGAACAACAACTCGGAAACCTGACCCACAAAGCTGTAGAGATGCTGCGCA 1171
 Db 1457 uThrArgAlaAsnValLysPheValGlnAspThr-----Se 1469
 QY 1172 GAATGAATGCTACATTTGGAGATACAGTCCG---CAGCGAGTGAAGATGTTAATGAA 1228
 Db 1469 rlystyrtlyrtyrlyspProaspiliserArgLuglnAlaIle-----AlaLeuLeuLys 1487
 QY 1229 AGAGAACAAAGATGCTACTTTTGGTCCGACAGCTGCTTACAAATGCAAGCAGAAC 1288
 Db 1487 sasprArgLuglnProglYAlaPheIleIleArgspSerHisSer----- 1501
 QY 1289 ATATGTTTTGGTGTGTTTATATGCAACAAAGCTCTACAAATGCAATCCGTTTCTCGA 1348
 Db 1502 -----PheArgGly-----AlaTyrGlyLeuAlaMetLysValAlaLase 1514

QY 1349 GAGCAATCAAGATTGGCCCTGGGACAGACTACGAGAAATGAGATGTTGATTCCT 1408
 DB 1514 rProPro-----ThValMetIndInsnLysLysGlyAspLleTh 1529
 QY 1409 GGAGACATCATTTGAACACTACACATATTTTCCATCTCTGCTATAGATGGAAAGACA 1468
 DB 1529 rAsnGluLeuValArgHisPhe-----LeuLleGluThrSerProAr 1543
 QY 1469 GCGTCGACGAGGAAACAGTGC----- 1490
 DB 1543 gGlyValLysLeuLysGlyCysProAsnGluProAsnPhgIlyCysLeuSerAlaLeuVa 1563
 QY 1491 -TACCTACACCCAGCCAGCCAGCTCTGCGCCAGGCTCTTCTCACTCACTGACCCAGCAGC 1549
 DB 1563 lTyrgInHisSerLleMetProLeuAlaLeu-ProCysLysLeuValLleProAspArgA 1583
 QY 1550 ACTTCATGATAGAAAGCCAGCCAGATATCCCGCATCATGAGCTGGCCCTGTCTCA 1609
 DB 1583 sPProMetGluGluLysLysAspAlaAlaSer----- 1593
 QY 1610 TTCCTGGCTCAATGATTCCTTCTTCATCTGATTTATTCGCAAGATATATTT 1669
 DB 1594 --ThrThrAsnSerAlaThrAspLeuLysGlnGlyAlaAlaCysAsnValLeuPheI 1613
 QY 1670 CTGTGCTTCAAGGATGATTTTGTACT 1698
 DB 1613 leAsnSerValGluMetGluSerLeuThr 1622

RESULT 10

A57075
 tensin - chicken (fragment)
 C.Species: Gallus gallus (Chicken)
 C.Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 21-Jul-2000
 C.Accession: A57075
 R.Chuang, J. Z.; Lin, D. C.; Lin, S.
 J. Cell Biol. 128, 1095-1109, 1995
 A.Title: Molecular cloning, expression, and mapping of the high affinity actin-capping d
 A.Reference number: A57075; MUID:95204530; PMID:7896874
 A.Accession: A57075
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-1792 <CHD>
 A.Cross-references: GB:I06662; NID:q212754; PIDN:AAA73949.1; PID:q212755
 A.Superfamily: SH2 homology
 F:1520-1629/Domain: SH2 homology <SH2>

Alignment Scores:

Pred. No.: 0.0178 Length: 1792
 Score: 140.00 Matches: 93
 Percent Similarity: 34.11% Conservative: 54
 Best Local Similarity: 21.58% Mismatches: 138
 Query Match: 4.58% Indels: 146
 DB: 2 Gaps: 21

US-09-856-061-1 (1-1721) x A57075 (1-1792)

QY 527 CGAGATCGGATAGCAGATACAGCTATTTCCAGATATGATGAGGCTCCCTCT 586
 DB 1347 ProGlyLeuGly-----ArgHisThrValSerProHisAlaProProGlySerProSer 1364
 QY 587 GTTACTCTCCAA-----GGCTTCTGTCTC 610
 DB 1365 LeuAlaArgHisGlnMetAlaAlaValProProGlySerProMetTyrgLysTySer-Se 1384
 QY 611 CACTGAGAGCAAAACCGAGATGAGATGACACACACTGGAAGAAATGCAAGCCTAC 670
 DB 1384 rProGluGluArg-----ArgProTh 1391
 QY 671 CTTCAGAGATGTCAGAGCCAAAGCTTTAAAGATTCAAATATACAAAATAAACAAGAC 730
 DB 1391 rLeuSer-----ArgInSerSerAlaSerGlyTyrgLn----- 1402
 QY 731 TCCTTTGCCACCTCTCGGCTGCTATCATCTTCCCAAGAATACCAACCC----- 782

DB 1403 -----ProProSerThrProSerPheProValSerProAlaTyrgLysProGlyThrSe 1420
 QY 783 -TTACCCCCAGCACACACAGAGAGAGAGCTGCTACTTC-----GCTCCAAA 829
 DB 1420 rThrProHisSerSerProAspSerAlaAlaTyrgInGlnLysSerProThrProG 1440
 QY 830 GCGCACCTTTCAGAAAGTCCAGAG-----GGGCCAGGAGAGAGAGCTGCAGAAAAGA 880
 DB 1440 nProAlaLeuProGluLysArgArgMetSerAlaGlyLysArgSerAsnSerLeuProAs 1460
 QY 881 CTTCAGTAGGCTCTTGAGACAGAAAGAAATCTCACCCACAGACAAAGCCAGAAATCTTC 940
 DB 1460 nTyrgAlaThrValAsnGly-----LysAlaSe 1469
 QY 941 TTGCCCA---TCATCAAAACCAAAACACAGAGAGAGTCCACTGCTGATTCAGCTCTTC 997
 DB 1469 rSerProLeuSerSerGlyMetSerSerProSerSerLysSerAlaAlaPheSerHi 1489
 QY 998 CTACATGCCA-----GGAAGCACAGTATACAGCCAGAGCCATACAGGTAGCATCA 1051
 DB 1489 sThrLeuProAspPheSerLysPheSerMet----- 1499
 QY 1052 GCACGTGCTCTGCTCAGAGATGCCAAGCTGCAGCCAGCCAGCCCTCGAATGCTCCCTA 1111
 DB 1500 -----ProAspLleSerProG 1505
 QY 1112 TGAANAACAAACTCGGANAACCTGACCCACAAAGCCTGATGAAAGATGCTGCGA 1171
 DB 1505 uThrArgAlaAsnValLysPheValGlnAspThr-----Se 1517
 QY 1172 GAATGATGTGTCATTTGAGAAATACAGTCGC---CAGCGAGTGAAGATGTTAATGAA 1228
 DB 1517 rLysTyrgLysPyrLysProAspLleSerArgAspGlnAlaIle-----AlaLeuLeuLys 1535
 QY 1229 AAGAAACAAGATGCTACTTTTGGTCCGACACTGCTTACAAAATCCAAGCAGCAAGCC 1288
 DB 1535 sAspArgGluProGlyAlaPheIleArgAspSerHisSer----- 1549
 QY 1289 ATATGTTTGGGTGCTTTTATGGGAACAAGGTGATGTAATGCTTCTCTCGA 1348
 DB 1550 -----PheArgGly-----AlaTyrgLysLeuAlaMetLysValAlaSe 1562
 QY 1349 GAGCAATCAACAGATTGGCCCTGGGACAGAGACTAGCAGGAAATGAGATTTGATTCGT 1408
 DB 1562 rProProPro-----ThValMetGlnInsnLysLysGlyAspLleTh 1577
 QY 1409 GGAACATCATTTGAACACTACACATATTTTCCATTCGCTAATGATGGGAAAGACA 1468
 DB 1577 rAsnGluLeuValArgHisPhe-----LeuLleGluThrSerProAr 1591
 QY 1469 GCGTCGACGAGGAAACAGTGC----- 1490
 DB 1591 gGlyValLysLeuLysGlyCysProAsnGluProAsnPhgIlyCysLeuSerAlaLeuVa 1611
 QY 1491 -TACCTACACCCAGCCAGCTGCTGCGCCAGGCTCTTCTCACTCACTGACCCAGCAGC 1549
 DB 1611 lTyrgInHisSerLleMetProLeuAlaLeu-ProCysLysLeuValLleProAspArgA 1631
 QY 1550 ACTTCATGATAGAAAGCCAGCCAGATATCCCGCATCATGAGCTGGCCCTGTCTCA 1609
 DB 1631 sPProMetGluGluLysLysAspAlaAlaSer----- 1641
 QY 1610 TTCCTGGCTCAATGATTCCTTCTTCATCTGATTTATTCGCAAGATATATTT 1669
 DB 1642 --ThrThrAsnSerAlaThrAspLeuLysGlnGlyAlaAlaCysAsnValLeuPheI 1661
 QY 1670 CTGTGCTTCAAGGATGATTTTGTACT 1698
 DB 1661 leAsnSerValGluMetGluSerLeuThr 1670

RESULT 11

S17983

[illegible]

Db	768	ssertysgluProValAlaValSerLysLysLysSerProProleuProth	788
QY	671	C-----TTCAG-----GATGTCAAGACCCAAACGCTTAAAGATTCAATACAC	715
Db	788	TValAspPheLysLIeArgLeuProValLThrAsnLIyAsnSerSerLysThrAlaSerP	808
QY	716	AAAAATAAACAGACACTCTTTGCCACCTCCCT---CGGGCTCTCTACCTCTCCCAAGAA	772
Db	808	olysLIeGluLysProLeuMetProProPAlaLysProPromeLLeuAlaProArgLy	828
QY	773	GTACCAACCCCTTA-----CCCCAGACACACAGAGAGAGACAGTCA--	815
Db	828	slenGluProSerAlaGlnPheAlaProProProSerProLLeIshIshIshIaGlyVa	848
QY	816	----TACTTGCGTCCA-----AAGCCACCTTTCC	841
Db	848	IGlnMetSerAlaProGlyAsnArgThrProIleAlaLysArgTyrIleProLleLeuP	868
QY	842	AGAAATCCAGAGG---GGGCCAGGACAGAGAGTGAAGAAAGACTTCAGTGGGCTCT--	896
Db	868	olysAlaSerArgProAsnProPheAlaAsnLIleProAsnAspValAsnArgLeuLeuLy	888
QY	897	-----GAGACAGAAAGAAATCTTCACCA	919
Db	888	sAspAlaGlyThrGluLIeLysSerLIeGlyGlyLysValGluAsnAsnSer	908
QY	920	CCAGACAAAGCCAGATCTCTTGCCCTCATCTCAACAAACAAACACAGAGAGTCCAC	979
Db	908	ralaGlnLysProHISLeuTyrGlyProLysGly-----GluThrLysMetGlyProPr	926
QY	980	TGCGATTGCCAGCTCTTCTCTCATGTCAGGAAGAACACAGATATCAAGACGACCATAC	1035
Db	926	oAlaLeuProAlaThrThrProSerGlnGlyAsnLysAlaGlyGlyAlaGlyAs	946
QY	1040	AGGTGATCGACGACCTGCTCTCT	1061
Db	946	nleuProMetSerAlaProProAsnLysGlyAsnSerSerAsnTyrLeuAsnLeuAl	966
QY	1062	-----GCTCAGAGATGCCAAGCTGCAGCGACGACAGCCCTCGAATCTGCTCTCA	1111
Db	966	aleuPheAsnSerAsnLysCysLysGlyLysGluAlaProProGlyCysArgThrProme	986
QY	1112	TCAAAAACAAACCTCGGGAACACTGCACCCACAAAGCTGATGACAGAGATGTCTGCA	1171
Db	986	tLysThrProAsnSerProIleTyrSerProSerSerProGln-----	1000
QY	1172	GAATGATGTGACATTTGAGATATCACT	1199
Db	1001	-----TyrValProSerTyrAsn	1006
RESULT	12		
	127877		
		hypothetical protein ZK470.5 - Caenorhabditis elegans	
		C.Species: Caenorhabditis elegans	
		C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999	
		A.Accession: T27877	
		R.Minix, P.	
		submitted to the EMBL Data Library, October 1995	
		A.Description: The sequence of C. elegans cosmid ZK470.	
		A.Reference number: Z20433	
		A.Accession: T27877	
		A.Status: preliminary; translated from GB/EMBL/DBJ	
		A.Molecule type: DNA	
		A.Residues: 1-443 <MIN>	
		A.Cross-references: EMBL:U039651; PIDN:AAA80397.1; CESP:ZK470.5	
		C.Genetics:	
		A.Gene: CESP:ZK470.5	
		A.Introns: 39/3; 83/3; 109/3; 214/3; 270/3; 332/3; 352/3; 416/2	
Alignment Scores:			
	0.018	Length:	443
Pred. NO.:	138.50	Matches:	46
Score:			

Db 379 ysleuilelysilpenehisargaspGlytyrGlypheserAspProLeu----- 396

QY 1390 ATGAGATGTTTCATCTCTGAGAGACATCATGAACACTAC 1430
 |||::||| |||
 ::::||| |||

Db 397 -----ThrpheAsnSerValValGIuLeuIleAsnHisTyr 408

Search completed: April 21, 2003, 12:32:00
 Job time : 89.7435 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:11 ; Search time 24.7582 Seconds

(without alignments)
5766.221 Million cell updates/sec

Title:

US-09-856-061-1

Perfect score:

3057

Sequence: 1 acgagcgcaactgcacagc.....taaaaaaaaaaaaaaaaaa 1721

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2p.model -DEV=xlp
-O=/cgn2.1/USPTO-spool/US09856061/runat_21042003_113018_3735/app.query.fasta_1.3150
-DB=SwissProt_40 -OPMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061@cgn1.1.65.@runat_21042003_113018_3735 -NCPu=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -LONLOG -DEV TIMEOUT=120
-MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412.5	13.5	533	1	LCP2_MOUSE
2	398	13.0	533	1	LCP2_MOUSE
3	148	4.8	722	1	P85B_HUMAN
4	146.5	4.8	217	1	GRAP_HUMAN
5	146.5	4.8	724	1	P85B_BOVIN
6	146	4.8	545	1	FGR_FSVGR
7	146	4.8	728	1	P85B_HUMAN
8	140.5	4.6	540	1	SCR_HUMAN
9	140	4.6	1744	1	TENS_CHICK
10	139.5	4.6	1603	1	PSC_DROME
11	137.5	4.5	592	1	ABP1_YEAST
12	137.5	4.5	722	1	P85B_MOUSE
13	137	4.5	724	1	P85A_BOVIN
14	136.5	4.5	561	1	3BP2_HUMAN
15	135	4.4	3421	1	TEGO_HSVB
16	133	4.4	675	1	BMX_HUMAN
17	133	4.4	724	1	P85A_HUMAN
18	133	4.4	724	1	P85A_MOUSE

19	133	4.4	724	1	P85A_RAT	063787	rattus norv
20	132.5	4.3	271	1	CRP_DROME	09xym0	drosophila
21	132.5	4.3	559	1	3BP2_MOUSE	006649	mus musculu
22	132	4.3	583	1	SHC_HUMAN	P29353	homo sapien
23	129	4.2	461	1	P55G_HUMAN	092569	homo sapien
24	129	4.2	751	1	FPI_MYTG	027409	mytilus gal
25	128.5	4.2	901	1	T298_HUMAN	015016	homo sapien
26	128	4.2	705	1	TRDN_RABIT	028820	cricetolagus
27	127	4.2	705	1	P55G_BOVIN	064604	bos taurus
28	127	4.2	1362	1	BRD4_HUMAN	060885	homo sapien
29	127	4.2	1386	1	ZAP3_MOUSE	093017	mus musculu
30	126.5	4.1	1271	1	BCR_HUMAN	P11274	homo sapien
31	125.5	4.1	2774	1	MAPA_RAT	P34926	rattus norv
32	125	4.1	962	1	P55G_MOUSE	064143	mus musculu
33	125	4.1	962	1	Y4E_SCHPO	009731	schizosacch
34	125	4.1	1210	1	AR4_HUMAN	P51825	homo sapien
35	124.5	4.1	585	1	PTNB_MOUSE	P35235	mus musculu
36	124.5	4.1	593	1	PTNB_RAT	006124	homo sapien
37	124.5	4.1	593	1	PTNB_RAT	P41499	rattus norv
38	124.5	4.1	1490	1	CRK7_HUMAN	Q9nyv4	homo sapien
39	123.5	4.0	845	1	CSW_DROME	P29349	drosophila
40	123.5	4.0	1736	1	ZOI_HUMAN	007157	homo sapien
41	122.5	4.0	875	1	FPI_MYTD	Q25460	mytilus edu
42	121.5	4.0	1147	1	KIN2_YEAST	P13186	saccharomyc
43	121	4.0	1217	1	AR4_MOUSE	088573	mus musculu
44	121	4.0	2442	1	CBP_HUMAN	092793	homo sapien
45	120.5	3.9	806	1	MK07_MOUSE	09wv58	mus musculu

ALIGNMENTS

RESULT 1

LCP2_MOUSE
ID LCP2_MOUSE STANDARD: PRT: 533 AA.

AC 060787;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell lymphoma;
RX MEDLINE=95221345; PubMed=7706237;
RA Jackman J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A., Koretzky G.A., Finkel P.R.;
RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells";
RT J. Biol. Chem. 270:7029-7032(1995).
CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND FYB.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES.
CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: U20159; AAC52189.1; --
CC HSSP: P12931; ISHD.

DR MGI:1321402; LCP2.
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00454; SAM; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS0001; SH2; 1.
 DR SH2 domain; Phosphorylation.
 KW DOMAIN 422
 FT 530
 SEQUENCE 533 AA: 60228 MW: 6DD877782C660C3 CRC64;

Alignment Scores:
 Pred. No.: 7,11e-21 Length: 533
 Score: 412.50 Matches: 148
 Percent Similarity: 39.88% Conservative: 51
 Best Local Similarity: 29.66% Mismatches: 175
 Query Match: 13.49% Indels: 125
 DB: 1 Gaps: 20

US-09-856-061-1 (1-1721) x LCP2_MOUSE (1-533)

OY 294 TTCGGTGTGAGATTCAGAAAGCTCTCTGCTG-----AAAAATAG 338
 DB 57 PheProLysLeuArgMetProLeuSerLysLeuSerGlnAspIleAsnLysGlu 76
 OY 339 TCATGAGCCAAAGCCTCAGCAGCCAAAGGCGGTGCGAGCGTTCGAAACCACTTCG 398
 DB 77 GluArgArgSerIlePheThrArgLysProGlnIleProArgPheLeuGlnLthrGlu 96
 OY 399 GATCAGACAGAAAGTGTGGCTGG-----422
 DB 97 SerHisGluGluAspGlyLysLysPheSerPheGluAspPtyrGluSerProAsn 116
 OY 433 -----GTCCAGCGGAGCAAAATGCAACAGTACCAAGCTACAGAGATCT-----470
 DB 117 AspAspAspProAspGlyGlu-----AspAspGlyAspPtyrGluSerProAsnGlu 133
 OY 471 GAGTTCACGCTGCTG-----AAGGATGGCCATCA-----500
 DB 134 GluGlnGlnIleLeuValAspAspAlaIleAspPtyrGluProProSerAsnAsnGlu 153
 OY 501 -----ATGAAATTTTACCAGCCAGCAGCTATCCAG-----GAATCGGAATAC 542
 DB 154 GluAlaLeuGlnSerSerIleLeuProProAsnSerPheHisAsnThrAsnSerMetLys 173
 OY 543 GCAGATACAGCCTATTTCAGGATATGATGAGGCTCCCTCTGTACT-----CCC 596
 DB 174 IleAspArgProProThrGlyLysValSerGlnGlnProProValProProLeuArgPro 193
 OY 597 AAGGCTCTGTC-----608
 DB 194 LysProAlaLeuProProLeuProThrGlyArgAsnHisSerProLeuSerProProHis 213
 OY 608 -----608
 DB 214 ProAsnHisGlnGluProSerArgSerGlyAsnAsnLysThrAlaLysLeuProAlaPro 233
 OY 609 TCACCTGAGACAAACCCAGGATGTGAGATGACACAGCTGGAAGAAGTGC-----662
 DB 234 SerIleAspArgSerThrLysProProLeuAspArgSerLeuAlaProLeuAspArgGlu 253
 OY 663 -----AAGCTTACCTTCAGAGATGTCAAGACCCCAACGCTTTAAAGA 704
 DB 254 ProPheIleLeuGlyLysLysProProPheSerAspLysProSerAlaProLeuGlyArg 273
 OY 705 TTCAATATACAAATAAATTAACAGATCTTGGCACCCTCTGCGCTGATACACTCTC 764
 DB 274 GluHisLeuProLysIleGlnLysProProLeuPro-----ProAlaMetAspArg 290
 OY 765 CCCAAGAAGTACCAACCTTACCCCGACACACACAGAGAGAGAGACTGATCTGCTGCT 824
 DB 291 HisGluArgAsnGlnArgLeuGlyProValThrThrArg-----303

OY 825 CCAAGCCACCTTCCAGAACTCCAGAGGGGCCCCAGAGAGAGTGCAGAAAGACTTC 884
 DB 304 ---LysProSerValProArgHisGlyArgGlyProAspArgGlu-----318
 OY 885 AGTAGGTCCTTGAGAGAGAAATCTCACCCACGACAGAAACCCAGAAATCTTCTTGC 944
 DB 319 -----AsnAspGluAspAspValHisGlnAlaGlyProLeuProGlnProSerLeu 334
 OY 945 CCATCATCAACCAACAGACAGAGATCCACCTGCGCATTCGACCTC-----995
 DB 335 ProSerMetSerSerAsnThrPheProSerArgSerValGlnProSerSerLysAsnThr 354
 OY 996 -----TCTACATCCAGAGAAAGCAGATATACAGCCAGACCATACAGTAGC 1046
 DB 355 PheProLeuAlaHisMetProGlyAlaPheSer-----GluSerAsnIleGlyPhe 371
 OY 1047 ATGACGACTGT-----CCTGCTCAGAGATGCCAAGCTGCAGCCAGCCAGCCCT---1097
 DB 372 GlnGlnSerAlaSerLeuProProTyrPheSerGlnGlyProGlyAsnArgProProLeu 391
 OY 1098 -----CGAATGCTCCCTATGAAACAAACATCGAGAAACCTGACCCACCA 1145
 DB 392 ArgSerGlnGlyArgAsnLeuProLeuProValProAsn-----ArgProGlnProPro 409
 OY 1146 AAGCCTGATGAGAGAGATGTC---TGCAGAAATGAAATGATGATTCGAGAAATACAGTCC 1202
 DB 410 SerProGlyGlnGluGlnLthrProLeuAspGlnLthrPtyrValSerTyrIleThrArg 429
 OY 1203 CAGCAGTGAAGATGTGTATGATGAAGACAGAAAGAGTGTACTTTTGTGTCGAGAC 1262
 DB 430 ProGlnAlaGlnAlaIleAlaLeuArgLysIleAsnLysPtyrIlePheLeuValArgAsp 449
 OY 1263 TGCTCTCAAAATCCAAAGCCAGAACCATATGTTTGTGTGTTTATGGCAACAGTCC 1322
 DB 450 SerSerLysLysThrAlaAsnAsnProTyrValLeuMetValLeuTyrLysAspLysVal 469
 OY 1323 TACAATGGAATACCGTTCCTGAGAGCAATCAACAGTTTGGCCGACAGCACTA 1382
 DB 470 TyrAsnIleGlnIleArgTyrGlnGlnGlnLeuSerGlnValLysLeuGlyThrGlyLeu 489
 OY 1383 CGAGAAATGAGATGTTGATCTGTGAGACATCATTCAGACATACATATATTTTCCC 1442
 DB 490 ArgGlyLysGlnAspPheLeuSerValSerAlaIleAsnLysPtyrPheArgLysMetPro 509
 OY 1443 ATCTGCTAATAGATGGGAAAGACAAAGCTGCAGCAGAGAAAGTCTACTTACC 1499
 DB 510 LeuLeuLeuIleAspGlyLysAsnArgLysSer---ArgTyrGlnCysThrLeuThr 527

RESULT 2
 LCP2_HUMAN
 ID LCP2_HUMAN STANDARD: PRT: 533 AA.
 AC 013094:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte
 DE protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
 GN LCP2.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Leukemia;
 RX MEDLINE=95221345; Pubmed=7706237;
 RA Jackson J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peliz G.A.,
 RA Koretzky G.A., Finkel P.R.;
 RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein
 RT associated with Gb2 in T cells.";
 RL J. Biol. Chem. 270:7029-7032(1995).
 J2]

RP SEQUENCE FROM N.A.
 RC Tissue-Prostate;
 RA Strauberg R.;
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
 CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND Fyb.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND
 CC PERIPHERAL BLOOD LEUKOCYTES. HIGHLY EXPRESSED ALSO IN T CELL AND
 CC MONOCYTIC CELL LINES, EXPRESSED AT LOWER LEVEL IN B CELL LINES.
 CC NOT DETECTED IN FIBROBLAST OR NEUROBLASTOMA CELL LINES.
 CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION BY Zap-70.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- DATABASE: NAME-PROT; NOTE-PROT 1:1-5(2000).
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1118450040_g.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U20158; AAC50135.1; -
 DR EMBL: BC016618; AAH16618.1; -
 DR HSSP: P12931; 1SHD.
 DR Genew; HGNC:6529; LCP2.
 DR MIM: 601603; -
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR Prodom: PD000093; SH2; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS50001; SH2; 1.
 KW SH2 domain; Phosphorylation.
 FT DOMAIN 133 136 POLY-GLU.
 FT DOMAIN 198 201 POLY-PRO.
 FT DOMAIN 422 530 SH2.
 SQ SEQUENCE 533 AA; 60188 MW; C5D22P31D36200C8 CRC64;
 Alignment Scores:
 Pred. No.: 6,99e-20 Length: 533
 Score: 398.00 Matches: 124
 Percent Similarity: 42.93% Conservative: 52
 Best Local Similarity: 30.24% Mismatches: 146
 Query Match: 13.02% Indels: 88
 DB: 1 Gaps: 14
 US-09-856-061-1 (1-1721) x LCP2_HUMAN (1-533)
 QY 456 GACTAGCAAGATCTGAGTTCACGCTGAGAGGCAATGCCATCAATGAAATTTTACCA 515
 DB 144 AsptylGluProProProSerAsnAspGluAlaLeuGlnAsn---SerIleLeuPro 162
 QY 516 GCGAGCCATCCAGCAATCGGAA-----TAGCAGATACACCGCATTTTCCAGATATG 569
 DB 163 AlalysProPheProAsnSerAsnSerMetTyrIleAspArgProProSerGlyLysThr 182
 QY 570 ATGAGAGCTCC-----CTTCTGTACTCCCAAGGCT 602
 DB 183 ProGlnGlnProProValProProGlnArgProMetAlaAlaLeuProProProAla 202
 QY 602 ----- 602
 DB 203 GlyArgAsnHisSerProLeuProProGlnThrAsnHisGluGluProSerArgSer 222
 QY 603 -----TCGCTCCACGAGAGCAACACGAGATG 635
 DB 223 ArgAsnHisLysThrAlaLysLeuProValProSerIleAspArgSerThrLysProPro 242
 QY 636 AGGATGACACAGCTGGAGAAAGTGGAC-----AAGCTACC 671

DB 243 LeuAspArgSerLeuAlaProPheAspArgGluProPheThrLeuGlyLysLysProPro 262
 QY 672 TTCAGAGT-----GTGAGAACCCAGCGCTTTAAAGATTTAAATACACAAATA 722
 DB 263 PheSerAspLysProSerIleProAlaGlyArgSerLeuGlyGlnHisLeuProLysIle 282
 QY 723 AACAGATCCCTTTCGACCTCTCCGCTGCTATCACTCTCCCAAGATACCAACC 782
 DB 283 GlnLysProProLeuProPro-----ThrThrGlnArgHisGluArg 296
 QY 783 TTACCCCAAGCACCACAG 842
 DB 297 SerSerProLeuPro-----GlyLysLysProProValPro 308
 QY 843 GAAGTCCAGAGGGGGCCAGCAGAGAGAGTGCAGAAAGACTTCAGTAGGGTCTTGAGACA 902
 DB 309 LysHisGlyTyrGlyProAspArgArgGlu-----AsnAsp 320
 QY 903 GAGAGAAATCTCACCACAGCAAGACAGAA-----TCCTCTGC 944
 DB 321 GluAspAspValHisGlnArgProLeuProGlnProAlaLeuLeuProMetSerSerAsn 340
 QY 945 CCATCATCAAAACCAACACACAGAGAGTCCAGCTCCAGCTCTCTACATG 1004
 DB 341 ThrPheProSerArgSerThrLysProSerProMetAsnProLeuProSerSerHisMet 360
 QY 1005 CCAGGAAAGCAGATATACAGCCACAGACCCATACAGGTAGATGACGATGCTCTGCT 1064
 DB 361 ProGlyAlaPheSer---GluSerAsnSerSerPheProGlnSerAlaSerLeuProPro 379
 QY 1065 CAGAGATGCCAGCTGCAGCCAGCCAGCCAGCT-----CGAATGCTGCC 1109
 DB 380 TyrPheSerGlnGlyProSerAsnArgProProIleArgAlaGluGlyArgAsnPhePro 399
 QY 1110 TATGAACACAAACATCGGAGAACCTGCAGCCCAAGCCCTGATGAGAGATGCTGG 1169
 DB 400 LeuProLeuProAsnLysProArgPro---ProSerProAlaGluGluLysAsnSerLeu 418
 QY 1170 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
 DB 419 AsnGluGluTyrTyrValSerTyrIleThrArgProGluAlaLeuAlaLeuArgLys 438
 QY 1230 GAGAACAGAGATGATCTTTTGTGTCGAGAGCTGTACAAATTCAGAGCAGACACA 1289
 DB 439 IleAsnGlnAspLysThrPheLeuValArgAspSerSerLysLysThrThrAsnPro 458
 QY 1290 TATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1349
 DB 459 TyrValLeuMetValLeuTyrLysAspLysValTyrAsnIleGlnIleArgTyrGlnLys 478
 QY 1350 AGCAATCAACAGTTTCCCTGGGACAGAGATACAGAGAAATGAGATGTTTGAATCTGTG 1409
 DB 479 GluSerGluValTyrLeuLeuGlyThrGlyLeuArgGlyLysGlnAspPheLeuSerVal 498
 QY 1410 GAAGACATCATTCAGCATACATATTTTCCATCTCTCTATAGATGGAAGACAAG 1469
 DB 499 SerAspIleIleAspTyrThrArgLysMetProLeuLeuLeuIleAspLysAsnArg 518
 QY 1470 GGTGACGAGCAAGACAGTCTACCTACC 1499
 DB 519 GlySer---ArgTyrGlnCysThrLeuThr 527
 RESULT 3
 P85B_RAT
 ID P85B_RAT STANDARD; PRT; 722 AA.
 AC 063788;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) (PcGlns-3-kinase p85-beta).
 GN PIK3R2.

D	b	55	GLYTPR-METPRGGLYPHEASNGUARGTHTARGINARGLYASPRHEPROGLYTHRTY	74
Q	y	471	----GAGTTC-----CAGCTGTCGAGGCGATGGCCA-----TCAAATGAA	505
D	b	74	rValGIuPHeuLeuGIyProValAlAlaLeuAlaArgProGIyProArgProArgGLyProAr	94
Q	y	506	AATTTTACACGACGCACCTATCCAG---GAATGGAAATACGCAATTCAC---CGCATATT	559
D	b	94	gProLeuProAlaArgProLeuAlaArgGLyProSerGluSerGlyNH1AsThrLeuAlaSerLe	114
Q	y	560	CCAGATATGATGAGGAGCTCCCTTCCTGTACTCC-----	596
			:::	
D	b	114	uAlaGIuGIuPHeSerProProGIuSerAlaProProIleuValLysLeuIleGIuAl	134
Q	y	597	-----AAGGTTCTGTCTCCACTGAGAGCAACACAG-----	629
D	b	134	AlleGIuGIuAlaGIuLeuAspSerGIuPHeTySerArgProGIuLeuProAlaProAr	154
Q	y	630	----GATGTGAGGATGACACACCTGTGAGAAAGTGTGACAAACCTTCACAGATGTGAG	685
D	b	154	gPhrAspTrpSerLeuSerAspLeuGIuGIuIntPrAspArgThrThrLeuTyTrsPala--	173
Q	y	688	AAGCCAGGCTTTAAAGGATTCAATTCACAAAATAAACAAAGCTCTTGGCACCTCC	745
D	b	174	-----ValLysGIyPHeu-----LeuAlaLeuPr	182
Q	y	746	TCGGCTGCTATCACTCTCCCAAGATACCAACCTTACCCCAAGACACACAGAGA	805
D	b	182	oAlaIaValValThr-----ProGIuAlaAl	191
Q	y	806	GAGCAGTCGATACCTTCGCTCCAAAG-----	830
D	b	191	AserIuAlaTyTrgAlaValArgIuValThrGIyProValGIyLeuValIleuGIuPr	211
Q	y	831	---CCACACTTTCACAAAGTCCAGAGGGGGCCAGACAGAGAGAGTCAAAAGACTTCAGT	889
D	b	211	oProThrLeuProLeuNH1SGInAlaLeuThrLeuAlaArgPheLeuLeuGIuNH1LysGIuAr	231
Q	y	880	GGTCTTGGAGCAGAGAGAGANTCTACACACAGACAAAGCAGAACTTCTTGGCCCATC	949
D	b	231	gVal-----	232
Q	y	950	ATCAACACAAACACACAGAAAGTCCA---CGTGCATTCGC-----AGCTC	994
D	b	233	-----AlaArgAlaArgAlaProSerProAlaThrAlaValNH1AlaLeuAlaLe	248
Q	y	995	TTCCTACATGCCA-----GGAAGACACAGTAT	1021
D	b	248	rAlaPheGIyProLeuLeuAlaArgAlaProProProArgIyGIuGIuLysArgIySerI	268
Q	y	1022	ACAAAGCCAGAGACCATACAGTAGACATGACAGCACTGCTCCTGACAGATGCCAAGCTCC	1081
D	b	268	uProAlaProAspPheProValLeuLeuGIuArgLeuValGIuGIuNH1LysValAspI	288
Q	y	1082	AGCCAGACACACCCCTGATGCTGCCGCTGTGAAACACAAATCTGAGAAACCTGAGCC	1141
D	b	288	uGIaAspThrAlaArgProAlaLeuPro---ProLysProSerLysValLysProAlaAr	307
Q	y	1142	CACA-----AAGCTGTAGAAAGATGCTGTGCGAAGTAATG	1180
D	b	307	oThrAlaLeuAlaNH1LysSerThrTrpProSerLeuAspAla-----GIuTr	324
Q	y	1181	GTACATTGGAGAAATACGTGCGCAGGAGGTGAGAAAGTGTAAATGAAAGAACACAGA	1240
D	b	324	rTyTrpGIyAspIleSerTrgGIuGIuValNH1ArgLeu-----ArgAspThrProAs	343
Q	y	1241	TGGTATTTTTGGTCCGAGACGTCTTCAAAATCCGTTTCTTCGAGACGATTCACAA	1300
D	b	343	rGIyThrPheLeuValArgAspAlaSerSerLysIleGIuGIuGIu---TyThrLeuTh	362
Q	y	1301	GGGTTTTATAGGCAACAAGTCTACATATGTGAATAATCCGTTTCTTCGAGACGATTCACAA	1360
D	b	362	rLeuAlaLysGIyGIyAsn-----AsnLysLeuIleLysValrPheNH1AspArgGIuNH1	380

OY 1361 GTTGGCCCTGGCGACAGACTACGAGAAATGAGATGTTGATCTGTGGAACATCAT 1420
 Db 380 STYGLYPHESERGLUPROLEU-----ThPhcyservalValGluLeuII 396
 OY 1421 TGAACACTACACATAT 1436
 Db 396 ESERHSTYARHIS 401

RESULT 4

GRAP_HUMAN STANDARD; PRT; 217 AA.
 ID GRAP_HUMAN 013588:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GRB2-related adaptor protein.
 GN GRAP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP MEDLINE=96218119; PubMed=8647802;
 RX Feng G.-S., Ouyang Y.-B., Hu D.-P., Shi Z.-Q., Gentz R., Ni J.,
 RA "Grp is a novel SH3-SH2-SH3 adaptor protein that couples tyrosine
 RT kinases to the Ras pathway."
 RL J. Biol. Chem. 271.12129-12132(1996).

CC -1- FUNCTION: COUPLE SIGNALS FROM RECEPTOR AND CYTOPLASMIC TYROSINE
 CC KINASES TO THE RAS SIGNALING PATHWAY.
 CC -1- SUBUNIT: ASSOCIATES THROUGH ITS SH2 DOMAIN WITH LIGAND-ACTIVATED
 CC RECEPTORS FOR STEM CELL FACTOR (KIT) AND ERYTHROPOIETIN (EPO).
 CC ALSO FORMS A STABLE COMPLEX WITH THE BCR-ABL ONCOPROTEIN. GRAP IS
 CC ASSOCIATED WITH THE RAS GUANINE NUCLEOTIDE EXCHANGE FACTOR SOS1,
 CC PRIMARILY THROUGH ITS N-TERMINAL SH3 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE GRB2 / SEM-5 / DRK FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; U52518; AAC50541.1; -
 DR HSSP; Q60631; IGRO.
 DR Genew; HGNC:4562; GRAP.
 DR MIM; 604330; -
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 2.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PRODOM; P0000066; SH3; 2.
 DR PRODOM; P000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 2.
 DR PROSITE; PS00001; SH2; 1.
 DR PROSITE; PS00002; SH3; 2.
 KW SH2 domain; SH3 domain; Repeat.
 FT DOMAIN 1 58 SH3 1.
 FT DOMAIN 60 152 SH2 2.
 FT DOMAIN 156 215 SH3 2.
 SQ SEQUENCE 217 AA; 25336 MM; 09PEC2F3BACDFAF8 CRC64;.

Alignment Scores: 0.0101 Length: 217
 Pred. No.: 146.50 Matches: 49
 Score: 146.50 Conservative: 29
 Percent Similarity: 44.83%

Best Local Similarity: 28.16% Mismatches: 48
 Query Match: 4.79% Indels: 48
 DB: 1 Gaps: 8

US-09-856-061-1 (1-1721) x GRAP_HUMAN (1-217)

OY 1131 AAACCTGACCCACAAAGCCTGATGAGAGATGTCTGCGAATGATGATCATTTGA 1190
 Db 56 LysProHisPro-----TTPyrSerGly 63

OY 1191 GAATACAGTCGCCAGGAGATGATGTGTAATGAAGAACAAAGATGCTACTTTT 1250
 Db 64 ArgIleSerArgInLeuAlaGluGluIleuMetLysArgSnhIleuGlyAlaPhe 83

OY 1251 TTGTCGCGAAGCTCTTACAAATCCAGGACGACACCATATGTTTGGTGTTTAT 1310
 Db 84 LeuIleArgGlu-----SerGluSerSerProGlyLupSerValSerValAsnTyr 101

OY 1311 GGGAAAGCTCTACATGTGCAAAATCCGTTTCCTCGAGCAATCAACAGTTTCCCTG 1370
 Db 102 GlyAspGlnValGlnHisPheLysValLeuArgGluAlaSerGlyLysTyrPheLeu--- 120

OY 1371 GGCACAGGACTACGAGAAATGAGATGTGATCTGTGGAACATCATTTGAACACTAC 1430
 Db 121 -----TTPglGluLulysPheAsnSerLeuAsnGluLeuValAspPheTyr 135

OY 1431 ACATATTTCCCATCTTCTGCTAATAGTGGAAACAAAGCTGCGACCGAGAAACAGTGC 1490
 Db 136 ArgThrThrThrIle-----AlaLysLysArgGlnIle 146

OY 1491 TACCTC-----ACCGACCCACGCTCCTC-----TCGCCAGGCTCTTCACACTCACTCTCC 1541
 Db 147 PheLeuArgAspGluGluIleProLeu-LeuLysSerProGlyLysAla---CysPheAlaGlnAl 165

OY 1542 AGCCAGGACTCTCATAGCTAAGAAAGCCAGCCACATATC----- 1581
 Db 165 aglnPheAspSerAlaGlnAspProSerGlnLeuSerPheArgGlyAspIleI 185

OY 1582 -----CCGATCAGTGGCTGGG 1599
 Db 185 egluValLeuGluIleArgProAspProHisTyrTrrArgGly 198

RESULT 5

P85B_BOVIN STANDARD; PRT; 724 AA.
 ID P85B_BOVIN
 AC P23726;

DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) (PtdIns-3-kinase p85-beta).
 GN PI3K2.

OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE=91191567; PubMed=1707345;
 RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Iarrea F., Panayotou G.,
 RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
 RA Courtneidge S.A., Parker P.J., Waterfield M.D.;
 RA "Characterization of two 85 kd proteins that associate with receptor
 RT tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase."
 RL Cell 65:91-104(1991).

RP CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.

RX MEDLINE=93049176; PubMed=1330535;

RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,
 RA Fry M.J., Blundell T.L., Mollner A., Waterfield M.D.;
 RA "Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
 RT domain with a PDGF receptor phosphorylation site: structural features

[illegible]

DE Tyrosine-protein kinase transforming protein fgr (EC 2.7.1.112).
GN V-FGR OR SRC-2.
OS Feline sarcoma virus (strain Gardner-Rasheed).
OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11775;
RN [1]
RX MEDLINE=84097512; PubMed=6318314;
RA Naharro G., Robbins K.C., Reddy E.P.;
RT "Gene product of v-figr onc: hybrid protein containing a portion of
RT actin and a tyrosine-specific protein kinase."; Science 223:63-66(1984).
RL Science 223:63-66(1984).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FGR
CC POLYPEPTIDE.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X00255; CAA25063.1; -
CC PIR: A00653; TVMVR.
CC HSSP: P00523; 2PTR.
CC
CC InterPro: IPR004001; Actin.
CC InterPro: IPR004000; Actin_like.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR000980; SH2.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00017; SH2; 1.
CC Pfam: PF00022; actin; 1.
CC Pfam: PF00069; pkinase; 1.
CC PRINTS: PR00401; SH2DOMAIN.
CC PRINTS: PR00109; TYRKINASE.
CC ProDom: PD000001; Euk_pkinase; 1.
CC ProDom: PD000093; SH2; 1.
CC SMART: SM00268; ACTIN; 1.
CC SMART: SM00252; SH2; 1.
CC SMART: SM00219; TYRK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00406; ACTINS_1; 1.
CC PROSITE: PS01132; ACTINS_ACT_LIKE; 1.
CC PROSITE: PS00001; SH2; 1.
CC PolyProtein: Tyrosine-protein kinase; Transferase; Phosphorylation;
KW ATP-binding; Oncogene; SH2 domain.
FT DOMAIN 23 157
FT DOMAIN 167 264
FT DOMAIN 286 539
FT NF_BIND 292 300
FT BINDING 314 314
FT ACT_SITE 405 405
FT MOD_RES 435 435
SO SEQUENCE 545 AA; 61516 MW; 7F4742EB7A7CA13E CRC64;

Alignment Scores:

Pred. No.: 0.0125 Length: 545
Score: 146.00 Matches: 69
Percent Similarity: 35.27% Conservative: 28
Best local Similarity: 25.09% Mismatches: 101
Query Match: 4.78% Indels: 78
DB: 1 Gaps: 11

US-09-856-061-1 (1-1721) x FGR_FSVGR (1-545)

QY 747 GGGCTGCTATCATCTCTCCCAAGAGTACCAACCTTACCCCGACACAGAGAG 806
|||||
DB 6 ATGPTAALAVAlCysArgProArgProLeuProProLeuProProThAlaMetLuGl 25
QY 807 AGCAGTGCATCTTC-----
|||
DB 26 GLUValAlaAlaLeuValIleAspAsnGlySerGlyMetCysLysAlaGlyPheAlaGly 45
QY 822 -----GCTCCAAAGCCCACTTTCAGAACTCCAGAGGGGGCCCA----- 861
|||||
DB 46 ASPAsPAlaProArgAlaValPheProSerIleVal-GlyAProArgHisGlnGlyVa 65
QY 862 -----GGCAGAGAGTGCACAAAGACTTCAGAGAGGCTCTTGAGCAGAGAAGAAT-- 912
||| |||
DB 65 IMetValGlyMetGlyGlnLysAspSerTyValGlyAspGlnAlaGlnSerLysArgI 85
QY 913 -----CTCACCCAGCAAAAGCCAGAA 935
|||||
DB 85 YIleLeuThrLeuLysTyProIleGlnHisGlyIleValIThrAsnTyPaspAspMetGl 105
QY 936 TCTTCTGCCCATCATCAACCAACCAACAGAGAGTCCACCTCCATTCGACGCTCT 995
|||
DB 105 uIysIleTyPHisHisThrPheTyAsnGlnLeuArgVal----- 118
QY 996 TCCTACATGCAGGAAGACAGTATACAGCCAGACCAATACAGTAGATGCAGCAC 1055
|||
DB 119 -----AlaProGlnGlu-----HisPr 124
QY 1056 TGTCTGCTCAGAGATGCCAGAGCTGACGAGCCAGCCCTCGAATG-----CT 1105
|||||
DB 124 oValleuLeuThrGluAlaProLeuAsnProLysAlaAsnArgGlnLysMetThrGlnI 144
QY 1106 GCCCTATGAAACACAAACCTCGGAGAACTGACCCCAAAAGCCGTGAGAGAGATGT 1165
|||
DB 144 eMetPheGluThrPheAsnIle-----ProSerAsnTyValAlaProValAspSerI 162
QY 1166 CTGCGCAATGAAATGTGTACATTGAGAAATACATGCCAGCGAGTGAA---GATGTGT 1222
|||
DB 162 eGlnAlaGlnGluIleTyPThrPheGlyLysIleGlyArgLysAspAlaGlnIleuGln 182
QY 1223 AATGAAGAGACAGAGAGATGCTTTTGGTCCAGAGCTGCTACAAATCCAAAGCC 1282
|||
DB 182 uSerProGlnLysAlaArgIylAlaPheLeuValArgGlnSerGluThrLysGlnAl 202
QY 1283 AGAACCATATGTTTGGTGTG-----TTTATGGAGACAGGCTACAA 1327
|||
DB 202 a-----TyrSerLeuSerIleArgAspIleAspGlnAlaArgGlyAspHisValLysH 220
QY 1328 TGTGAATCCGTTTCTCGAGAGCAATCAACAGTTTGGCCCTGGGACAGAGCTACGAG 1387
|||
DB 220 sTyTyLysIleArgLysLeuAspThrGly-----GlyTyTyThrIleThrTh 235
QY 1388 AAATGAGATGTTGATTCGTGTGAGACATCATTTGAACACTAC 1430
|||
DB 235 rArgAlaGlnPheAsnSerValGlnGlnLeuValGlnHisTy 249
|||
RESULT 7
P85B_HUMAN STANDARD: PRT; 728 AA.
ID P85B_HUMAN
AC 000459;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
DE p85-beta subunit) (PtdIns-3-kinase p85-beta).
GN PIK3R2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98241181; PubMed=9582025;

ID	SCK_HUMAN	STANDARD:	PRT:	540 AA.
AC	P98077; Q9NPJ5; O60230;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Protein SCK (Fragment).			
GN	SKC.			
OS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Db	155	ThpArpTprSerLeuSerAspValAspGlnTprsrphrAlaAlaLeuAlaAspArgLysLeu	174	
QY	573	GAGGCTCCCTCTTCG---TTACST-----CCAGGCTCTGTCTCCACT	614	
Db	175	LySserPheLeuAlaLeuAlaLeuAlaProLeuValThrProGlnAspSerAlaGlnAla	194	
QY	615	GAGACAAACACAGGATGTGAGATGACACACAGCTGGAGACAGCTGGACAAAGCTTCTTC	674	
Db	195	ArgAlaGlnAlaLeuAlaArgGlnAlaAlaGlnProValGlnProAlaLeuGlnProProThrLeu	214	
QY	675	AAGATGTGCAGAACCCAGCTTTAAAGATTCAATATACAAAAATTAACAGACTCT	734	
Db	215	ProLeuAlaThrAlaLeuThrLeuArg---PheLeuGlnAlaLeuGlnArgValAla	233	
QY	735	TTGGCACTCTCTCGG-----CGNGTATC---ACTCTCCCAAGAAATGACACCTTAA	785	
Db	234	ArgAlaGlnAlaProAlaLeuGlnProAlaValAlaGlnAlaLeuAlaThrPheGlnProLeu	253	
QY	786	-----CCSCAGACACACACAGAGAGAGAGATGCATCTCTCCCTCCAAAGCCAC	836	
Db	254	LeuLeuAlaGlnAlaProProGlnProGlnSerSerProProGlnGlnAlaProAlaProSer	273	
QY	837	TTTCCAGAAAGTCCAGAGGGGCCAGGACAGAGAGTGCACAAAGACTTCAAGTACGCTCT	896	
Db	274	GlnPro-----SerProAspPheProAlaLeuLeu	283	
QY	897	GGAGCAGAGAAAGATCTCACACACACAGACAAAGCCAGAAATCTCTCCAGCATACAAAC	956	
Db	284	ValGlnLysLeuLeuGlnGlnAlaLeuGlnGlnGlnGlnAlaAlaProGlnAlaLeuPro	303	
QY	957	CAAAACACACAGAAAGACTCCACCTGCCATTTGCCAGCTTTCTTACATGCCAGAAAGAC	1016	
Db	304	ProLysProProLysAlaLysProAla-----	312	
QY	1017	AGTATACAGACGACAGACATACAGATGATGATGACACACCTGTCTGTCTCAAGATGCCAA	1076	
Db	312	-----	312	
QY	1077	GCTGCAGCCAGCCACAGCCCTCGATGCTGCTCCCTATGAACACAAACCTGAGAGAACT	1138	
Db	313	-----ProThrValLeu-----AlaAsnGlnLysArgPro	322	
QY	1137	GACCCACAAAGCCTGATGAGAGAGATGCTGSCAGAAATGAATGTGATCTTGGAGAAATAC	1196	
Db	323	-----ProSerLeuGlnAspAla-----GlnTprLysTprLysPhe	335	
QY	1197	AGTGCACAGGACAGTGAAGATGTGTTAATGAAGAGACACAGATGATCTTTTGGTGC	1256	
Db	336	SerArgLysGlnLysValAsnGlnLysLeu---ArgAspThrProAspGlnThrPheLeuLeu	354	
QY	1257	CGAAGCTGCTCAAAATTCGACAGGACAGAAACATATGTTTGATGGTGTGCTTTAATGGAAC	1316	
Db	355	ArgAspAlaSerSerLysLysIleGlnGlnLys---TyrThrLeuThrLeuArgLysGlnLys	373	
QY	1317	AAGCTTACAAATGAAATCCCTTTCTCCGAGACACATCAACAGTCTTGCCCTGGGACACA	1376	
Db	374	Asn-----AsnLysLeuIleLysValPheAlaSerArgGlnLysIleGlnPheSerGln	391	
QY	1377	CGACTACGAGGAATAGATGTTGATTTGTTGTGGAGACATCATTTGAACSTACACATAT	1436	
Db	392	ProLeu-----ThrPheCysSerValAlaSerLeuIleAsnAlaIleLysIleArgHis	407	
RESULT 8				
SCK_HUMAN				
ID	SCK_HUMAN	STANDARD:	PRT:	540 AA.
AC	P98077; Q9NPJ5; O60230;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Protein SCK (Fragment).			
GN	SKC.			
OS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98175964; PubMed=9507002;
 RX Nakamura T., Murooka S., Sanokawa R., Mori N.;
 RT "N-Shc and Src, two neuronally expressed Shc adapter homologs. Their
 RT differential regional expression in the brain and roles in
 RT neurotrophin and Src signaling";
 RN J. Biol. Chem. 273:6960-6967(1998).
 RN [2]
 RP SEQUENCE OF 72-499 FROM N.A.
 RX MEDLINE=95090462; PubMed=7527937;
 RA Kavanaugh W.M., Williams L.T.;
 RT "An alternative to SH2 domains for binding tyrosine-phosphorylated
 RT proteins";
 RN Science 266:1862-1865(1994).
 RN [3]
 RP SEQUENCE OF 72-328 FROM N.A.
 RA Lamerda J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Brower A., Ramirez M.,
 RA Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A.,
 RA Gannes J., Dangnan L., Erler A., Christensen M., Georgescu A.,
 RA Avila J., Liu S., Altix C., Andreise T., Frankheim M.,
 RA Amico-Keller G., Coeffield J., Duarte S., Lucas S., Bruce R.,
 RA Thomas P., Quan G., Krommiller B., Arellano A., Sanders C., Ow D.,
 RA Nolan M., Trong S., Kobayashi A., Olsen A.S., Carraro A.V.;
 RT "Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a
 RT serine protease gene cluster";
 RT Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 149-540 FROM N.A.
 RA Carim L., Estivill X., Sumoy L., Escarceller M.;
 RL Submitted (JULY-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN LIVER, ALSO PRESENT IN BRAIN.
 CC -1- SIMILARITY: CONTAINS 1 PID DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: TO SHC TRANSFORMING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb.ch/announce/
 CC or send an email to license@isb.sib.ch).
 CC -----
 DR EMBL: AB001451; BAA25798.1; -;
 DR EMBL: AC006124; AAC97988.1; -;
 DR EMBL: AL360254; CAB96175.1; -;
 DR HSSP: P29353; 1MIL.
 DR MTM: 605217; -;
 DR InterPro: IPR000050; PID.domain.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00640; PID; 1.
 DR PRINTS: PRO0401; SH2DOMAIN.
 DR PRINTS: PRO0629; SHC2PIDOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00462; PTB; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS01179; PID; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR SH2 domain.
 FT NON_TER 1 1
 FT DOMAIN 105 287 PID.
 FT DOMAIN 445 536 SH2.
 FT CONFLICT 72 72 A -> P (IN REF. 2).
 SQ SEQUENCE 540 AA; 57623 MW; 0D83700910B1A5DB CRC64;

Alignment Scores: 0.0297 length: 540
 Pred. No: 140.50 Matches: 101
 Score:

	Percent Similarity:	34.59%	Conservative:	55
	Best Local Similarity:	22.39%	Mismatches:	196
	Query Match:	4.60%	Indels:	99
	DB:	1	Gaps:	19
US-09-856-061-1 (1-1721) x SCK_HUMAN (1-540)				
QY 243 GTCCCAAGGACCATGACGACGAGGCAATTAAGACAGCAAGAA-----				290
DB 148 VALPIROGVALARGIYserTrpIysLysValProAsnIValAlaLeuAlaSerVal				167
QY 291 ---GGATTGCGATGATGAGATTCAGAAAGCTCTCTGCTGCAAAAATAGTCATGGCCA				347
DB 168 LeuGlyLysSerAsnLeuArgPheAlaGlyMetSerIleSerIleHisIleSerThrAsp				187
QY 348 AGCCTCGACGAGGCAAGAGGGGGGTGCGAGGGTTCGATCGACCACTCCGATCCAGAGA				407
DB 188 GlyLeuSerLeuSerValProAlaThrArgGlnValIle-----AlaAsnHisHisMet				205
QY 408 AGGAACATGCTGGGTGCCAGGTGAGAGAAATGCAACAGTACAGCACTAC-----				461
DB 206 ProSerIleSerPheAlaSerGlyGlyAsp-----ThrAspMetThrAspTyrValAla				223
QY 462 -----GAAGATCCTGATGTTCCAG-----CTGCTGAAGCATGGCCATCA				500
DB 224 TyrValAlaLysAspProIleAsnGlnArgAlaCysHisIleLeuGlnCysGlyGly				243
QY 501 ATGAAATTTTACCAGCCAGACCTATCCAGCAATGC-----GAATACGCAATACAGCC				554
DB 244 Leu-----AlaGlnSerIleIleSerThrValGlyGlnAlaPheGlnLeuArg				259
QY 555 TATTTCCAGGATATGATGAGGCTCC---CTTCTGTACTCTCCAGGCTTGTCTCC				611
DB 260 PheLysGlnTyrLeuHisSerProPolsValAlaLeuProGlnArgLeuAlaGly				279
QY 612 ACTGACAGACAAACAGGATGTCAGATGACACACTGGAA-----				653
DB 280 ProGlnLysSerAlaTrpGlyAspGlnLysSerLeuGlnHisAsnTyrTrpAsnSer				299
QY 654 -----GAAGTGACAGGCTTACTCTTCAAGATGTCAGCAAGCAAGCTTTAA				701
DB 300 IleProGlyLysGlnProProLeuGlyGlyLeuValAspSerArgLeuAlaLeuTrgln				319
QY 702 GGATTCATATACAAATAATTAACAAGACTCTTCCACCTCCGCTGCTGATCACT				761
DB 320 ProCysAlaLeuThrAlaLeuAspGlnGlyProSerProSerLeuArgAspAlaCysSer				339
QY 762 CTCCCAAGAGTACCAACCTTACCCAGCACACACAGAGAGACATGATCTTC				821
DB 340 LeuProTrpAspValGlySerThrGlyThrAlaProProGlyAspGly-----				355
QY 822 GCTCAAGCCACCTTCCAGAAATCCAG-----AGGGGCCGACAGGAGAGT				872
DB 356 -----TyrValGlnAlaAspAlaArgGlyPro-----				364
QY 873 GCAAAAGACTTCAAGTAGGCTCTTGAGAGCAAGAAAGATCTGACCAAGCAAGCAAGCA				932
DB 365 ---ProAspHisGlnGlnHisLeuTyrValAsnThrGlnGlyLeuAspAlaProGlnPro				383
QY 933 GAATCTTCTTGCCCATCATCAAAACCAACACACAGAGAGTCCACTGCGATTCGCACG				992
DB 384 GluAspSer-----				386
QY 993 TCTTCTACATGCCAGGAAGACAGTATCAAGCCAGAGACATACAGGTAGCATGACAG				1052
DB 387 -----ProLysLysAspLeuPheAspMetArgProPheGlnAspAlaLeuLys				402
QY 1053 CACTGTCTGCTCAGAGATGCCAAGCTGACAGCCAGCAAGCCCTCGAATGCTGCTAT				1112
DB 403 -----LeuHisGlyLysSerValAlaGlyValThrAlaAlaProLeu				419
QY 1113 GAAACACAAACTCGGAGAAACCTGCAC-----CCCAACAAAGCTGATGAGAAAGAT				1163

Query Match: 4.56% Indels: 118
DB: 1 Gaps: 20
US-09-856-061-1 (1-1721) x PSC_DROME (1-1603)

QY 278 GACAAAGAAAGAAAGATTCGATGATTCGAAAGACGCTCTGCTGTAAGAAATAG 337
DB 613 AAPPProgluArg-----GuileVallyProleuLysProgluLysglu 628
QY 338 GTCATGGCCACCTCAGAG-----TGCAGAAAGCGCGTTCGAGCGTTCTGA 388
DB 629 SerArgSerLysLysLysAspLysAspLysSerProLysSerSerSerSerSer 648
QY 389 ACCACTTCGCGATCAGACAGAACTGAGCTGGCTGCC----- 427
DB 649 SerSerSerGlyGluArgLysArgLysSerProSerProLeuThrValProLeuThr 668
QY 428 ---AGGTGAGAAATATCAGATACAGACACTAC----- 461
DB 669 IleArg-ThrGluArgIleMetSerProSerGlyValSerThrLeuSerProArgValTh 688
QY 482 -----GAAGATCCCTGAGATTCAGCTGCTGAAGCATGCCA----- 497
DB 688 rSerGlyAlaPheSerGluAspProLysSerGluPheLeuSerPheAlaLeuLysPr 708
QY 498 -----TCAATGAATATTTACAGC 517
DB 708 OlleLysValLysValLysLysProGluArgThrLeuAsnAsnArgAlaIleThrProPr 728
QY 518 CAGACCT---ATCCAGAAATGGAATACGACATACAGC-----TATTTCCAGGAT-- 566
DB 728 oSerProSerValGlnGlnSerAlaSerProLysSerLysLysAsnLeuAspAspSe 748
QY 567 ---ATGATGAGAGGCTCCCTTGTGTTACCTCCAGGCTTCTGCTCCATGAGAGCA 622
DB 748 rIleLeuMetLysProProSerCysMetProProLysSerIleAlaSerLysArgLys 768
QY 623 AACCAAGGAT-----GTGAGAGATGACACAGCTGAGAAAGTGACAGAGCTTAC 670
DB 768 sSerLysGluProValLysAlaValSerLysLysGlnLysLeuSerProProLeuProTh 788
QY 671 C-----TTACAG-----GATGTCAGAGCCAAAGCCTTTAAAGGATTCAAATACAC 715
DB 788 rValAspPheLysIleArgLysProValThrAsnGlyAsnSerSerGlyThrAlaSerPr 808
QY 716 AAAAATAAACAAGACTCTTGGCCACCTCT---CGGCGCTGCTTCACCTGCCCAAGAA 772
DB 808 oLysIleGluLysProLeuMetProProProAlaLysProProMetLeuAlaProArgLys 828
QY 773 GTACCAACCTTA-----CCCCAGACACACAGAGAGAGAGAGTGA-- 815
DB 828 slenGlnProSerAlaGlnPheAlaProProProSerProIleHisIshAlaGlyVa 848
QY 816 ---TACTTGGCTTCA-----AAGCCACCTTTCC 841
DB 848 lGlnMetSerAlaProLysAsnArgThrProIleAlaLysArgGlyGlnProIleLeu 868
QY 842 AGAAGTCCAGAG---GGGCCAGAGAGAGAGTGAAGAAAGACTTCAGTGGGCTCT-- 896
DB 868 oLysAlaSerArgProAsnProPheAlaAsnIleProAsnAspValAsnArgLeuLys 888
QY 897 -----GGAGCAGAGAAAGAAATCTCACCA 919
DB 888 sAspAlaGlyThrGluIleLysSerIleGlyGlySerValGlnLysAsnSerAsnSe 908
QY 920 CCAGACAAACCCAGAACTCTTGGCCATATCAACCAAAACACACAGAGAGTCCACC 979
DB 908 rAlaGlnLysProHisLeuTyrgLysProLysGly-----GluThrSwetGlyProp 926
QY 980 TGCCATTCGACGCTCTTCATATGCGAGAAAGACAGATATCAAGCCAGAGACCATAC 1039
DB 926 oAlaLeuProAlaThrThrProSerGlnGlyAsnLysAsnValGlyLysGlnAlaLys 946

QY 1040 AGTAGCATGACGACACTGCTCT----- 1061
DB 946 nLeuProMetSerAlaProProAsnLysGlyAsnSerSerAsnAsnTyrlLeuAsnLeuAl 966
QY 1062 -----GCTCGAATATCCAAAGCTGACGCCAGCCAGACGCTTCGATGCTGCCCTTA 1111
DB 966 aLeuPheAsnSerAsnLysCysLysGlyLysGlnAlaProProGlyCysArgThrProMe 986
QY 1112 TGAAMAACAACCTCGAGAAAGCTGAGACCCACAAAGCCTGATGAGAAAGATGCTGGA 1171
DB 986 tTyThrProAsnSerProIleTySerProSerSerProLyn----- 1000
QY 1172 GAATGAATGTCATCTTGAGATACACT 1199
DB 1001 -----TyrlValProSerTyrlAsn 1006

RESULT 11
ABP1_YEAST
ID ABP1_YEAST STANDARD PRT: 592 AA.
AC P15891;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Actin binding protein.
GN ABP1 OR YCR088W OR YCR88W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
ON NCBI_TaxID=4932;
RX MEDLINE=90136906; PubMed=2405279;
RA Drubin D.G., Mulholland J., Zhu Z., Botstein D.;
RT "Homology of a yeast actin-binding protein to signal transduction
RN Nature 343:288-290(1990).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90136906; PubMed=2405279;
RA Drubin D.G., Mulholland J., Zhu Z., Botstein D.;
RT "Homology of a yeast actin-binding protein to signal transduction
RN Nature 343:288-290(1990).
RL [2]
RP SEQUENCE FROM N.A.
RA Frontali L., Grisanti P.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: MAY BE INVOLVED IN THE SPATIAL ORGANIZATION OF CELL
CC SURFACE GROWTH. AN OVERPRODUCTION OF ABP1 CAUSES THE ASSEMBLY OF
CC THE CORTICAL ACTIN SKELETON AT INAPPROPRIATE SITES ON THE CELL
CC SURFACE, RESULTING IN DELocalized SURFACE GROWTH.
CC - SUBCELLULAR LOCATION: CORTICAL CYTOSKELETON.
CC - SIMILARITY: TO SIGNAL TRANSDUCTION PROTEINS AND MYOSIN I.
CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
DR EMBL: X51780; CAA36075.1; -
DR EMBL: X59720; CAA42253.1; -
DR PIR: S19503; LBY.
DR HSSP: P29355; ISEM.
DR SGD: S0000684; ABP1.
DR InterPro: IPR002108; Actbind.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00241; cofilin_ADF; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00102; ADF; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
KW Cytoskeleton; Actin-binding; SH3 domain; Repeat.
FT SIMILAR 1 142 TO YEAST COFILIN.
FT DOMAIN 86 97 ACTIN-BINDING (POTENTIAL).

```

FT DOMAIN 532 592 SH3
FT REPEAT 200 575 3 X 10 AA APPROXIMATE REPEATS.
FT REPEAT 200 209 1
FT REPEAT 436 445 2
FT REPEAT 566 575 3
FT CONFLICT 58 58 L -> S (IN REF. 1).
FT CONFLICT 312 312 K -> I (IN REF. 1).
SO SEQUENCE 592 AA; 65576 MW; 3952351070494AA CRC64;

Alignment Scores:
Pred. No.: 0.0482 Length: 592
Score: 137.50 Matches: 84
Percent Similarity: 36.88% Conserved: 58
Best Local Similarity: 21.82% Mismatches: 135
Query Match: 4.50% Indels: 108
DB: Gaps: 20

US-09-856-061-1 (1-1721) x ABPI_YEAST (1-592)
QY 221 GAAGATCCCTCTCTTTGGCAGGTGCGCAGGACCATGACCGAGGCAATAAAGGAC 280
DB 244 GIUASPPHrgrleuval-GluysPro-----ThralAlaGlySerIlysIleas 260
QY 281 AACGAAAGAGATTCGATCTGAGATTCAGAACGTCCTCTGTAATAAATAGTTC 340
DB 260 ProSerSerAspIleAlaIleuLys-----AsnGluSerIlysLeuLysArgAspse 278
QY 341 ATGGCCAGCTCAGCAGTCCCAAAGGCGGTTCGAGCGGTTCGAAACCACTCCGGA 400
DB 278 r-----GluPhenSerPheLeuGlyThrIlyPro----- 289
QY 401 TCACAGAGAACTTGGCTGGGTCCAGGTGAGAAATAATGCAACAGTAAACAGCTA 460
DB 290 -----ProSerMetThrGluSerSerLeuLysAsnsp-- 300
QY 461 CGAAGATCCTGATCCAGCTGCTGGAAGCATG-----CCATCAATGAA 505
DB 301 -----AspAspIysValIleLysGlyPheArgAsnGluLysSerProIaGlnle 317
QY 506 AATTTTACCGCAGACCTTCCAGAGATGGAATACGCAATACGCAATACGCTATTCCGGA 565
DB 317 UTTPAlaGluArgLysAlaLysGlnAsnSerGlyAsnIaGluThrIys----- 333
QY 566 TATGATGAGGCTCCCTCTCTTACCTCCCAAGGCTTCTCTCCACTGAGAGACAAC 625
DB 334 -----AlaGluAla-----ProLysProGluValProLysasp-- 344
QY 626 CAGGATGTGAGATGACACAGCTGGAAGAAGTGGACACCTTCAAGATGTCTAG 685
DB 345 -----GluProGluGluGluProAspValLysAspLeuLys 356
QY 686 AGCAAGCGTTTAAAGGATTCAATACAAAAATAAACAGCTCTTG----- 737
DB 356 sSer--LysPheGluLysLeuAlaIleSerGluLysGluLysGluLysGluLysLeu 375
QY 738 -----CCACTCTCTCTG-----CCTGCTATCATCTCTCCCAAGAAGTAC--CA 778
DB 375 sPheAlaProProLysLysSerGluProThrIleIleSerProLysProPheSerLys 395
QY 779 ACCCTTACCCCGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838
DB 395 sPheGluGluProValLysAlaGluGluAlaGlu-----GlnProLysThrAspTy 412
QY 839 TCCAGATTCAGAGAGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 898
DB 412 rLysLysIleLys-----AsnProLeuProGlu 421
QY 899 AGCAGAGAGAGATTCACACACAGACAGAAACCAAGATCTTCT----- 941
DB 421 yMetHisIleGluLysAsnGluGluGluProGluGluLysAsnspAspAspAsp 441
QY 942 -----TGCACATCATCAAC----- 956

```

```

DB 441 pasgIuspgIuaIaIaGlnProProLeuProSerArgAsnValAlaSerGlyAlaPr 461
QY 957 -----CAAAACACACAGAGAGATTCACCTCGCATTCGACCTCTTACATCCAGAAA 1012
DB 461 oValGlnLysGluGluProGluGluGluGluGluGluGluGluGluGluGluGlu 479
QY 1013 GCACAGTATACAGACCGACGAGACCATACAGATGAGATGAGAGAGAGAGAGAGAG 1072
DB 479 gasnSerIleProAlaProLysGluGluGluGluGluGluGluGluGluGluGlu 499
QY 1073 CCAAGCTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1132
DB 499 eGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 519
QY 1133 ACCTGACCCCAAGAG-----CCTGATGAGAGAGATGCTCTGCAAGATGATGATG 1189
DB 519 oProProProAlaGluAlaThrProGluLysLysProLysGluAsnProIleAlaThrAl 539
QY 1190 AGAATACAGTCCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1249
DB 539 agIuTyraPtyrAspAlaIaGluLys-----AsnGluLeuThrPh 553
QY 1250 TTGGTCCGAGAC 1262
DB 553 eValGluAsnsp 557

RESULT 12
P85B_MOUSE
ID P85B_MOUSE STANDARD: PRT; 722 AA.
AC 008908;
DT 15-DEC-1998 (rel. 37, created)
DT 15-DEC-1998 (rel. 37, last sequence update)
DT 15-JUN-2002 (rel. 41, last annotation update)
DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
DE p85beta subunit) (Ptdins-3-kinase p85-beta).
GN PI3K2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=NIH SWISS;
RC MEDLINE=98241181; PubMed=9582025;
RA Janssen J.W.G., Schleithoff L., Bartram C.R., Schulz A.S.;
RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase
RT p85beta subunit and HMOX8, a putative deubiquitinating enzyme.";
RL Oncogene 16:1767-1772(1998).
CC -!- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC -!- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -!- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -!- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: Y13569; CAJ73903.1;
DB HSSP: P23727; 2PNB.
DB MGD: MGI:1098772; PIK3R2.
DB InterPro: IPR001720; PI3kinase_P85.
DB InterPro: IPR000198; RhoGAP.
DB InterPro: IPR000980; SH2.
DB InterPro: IPR001452; SH3.

```


DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PIKINSEPS.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD00066; SH3; 1.
 DR PRODOM: PD00093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 KW SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 125 257 SH3.
 FT DOMAIN 324 419 SH2 1.
 FT DOMAIN 616 710 SH2 2.
 SQ SEQUENCE 722 AA; 81251 MW; AB4D49AF30CEC567 CRC64;

Alignment Scores:
 Pred. No.: 0.0495 Length: 722
 Score: 137.50 Matches: 102
 Percent Similarity: 36.498 Conservative: 52
 Best Local Similarity: 24.178 Mismatches: 151
 Query Match: 4.50% Indels: 117
 DB: 1 Gaps: 25

US-09-856-061-1 (1-1721) x P85B_MOUSE (1-722)

QY 365 AGGCGGCTGTCAGCGGCTTTCGAAACCACTCCGATCCAGACAG-----GAACTT 415
 DB 35 ATGVALAALALEUGLALALEUGLYVALAALASPGLYGLIARGLYSPROHISANVAL 54
 QY 416 GCGTGGGCTCCAGCT---GGAGAAATAATGCAAGTAACAAGCTCGAAGATCC-- 470
 DB 55 GLYTRP-METPROGLYPHEASNGLIARGLYTHARGLINARGLYASPPHEPROGLYTHRY 74
 QY 471 ---GAGTTC-----CACTGCTGAAGGATGCGCA-----TCAATGA 505
 DB 74 TVALGLIUPHEUGLYPROVALAALALEUALARPROGLYPROALYPROALYPROAR 94
 QY 506 AATTATACGAGCAGATCCATCCAGATCCGAAATACCAATACCACTTATTCACAGA 565
 DB 94 GPROLEUPROALATARGPROLEUASPGLYSERSEGLYSEGLY---HSLLEULEUPROAS 113
 QY 566 TATGATGAG-----GCTCCCTTCTGTACTCTCC----- 596
 DB 113 PLEUALGLIUPHESETPROPROALAPROPROALAPROPROLEUVALYLSLEUVALGL 133
 QY 597 -----AAGCTTCTGTCTCCACTGAGAGACAAACAGG----- 629
 DB 133 UALALEGLIUNLALAGLULEUASPSERGLUCYSEYSEGLYSPROGLULEUPROALATH 153
 QY 630 -----GATGTAGAGATGACAGACAGAGAGAGTGAAGAGCTTCAAGATGT 682
 DB 153 FARGTHASPRIPRSEULEUSERASPLEUGLINTPRSPARGTHRALALEUTRYASPA 173
 QY 683 CAGAACCCACGCTTAAAGATTCAATACACAAATAAACAAGACTCTCTTGCCACC 742
 DB 173 A-----VALYSGLYPHELEU-----LEUALALE 181
 QY 743 TCCTGGGCTCTATCACTCTCCCCAGAGATACCAACCTTACCCAGACACAGAGA 802
 DB 181 UPROALALALAVALLATH-----PROGLUAL 190
 QY 803 GGAGAGCAGTGCATATCTGCTCCAAAGCCACCTTTCAGAGAGTCCAGAGGGGCCAG 862
 DB 190 AALALAGLULATYARGALALEUAG-----GLIUALALAGLY----- 203
 QY 863 GCAGAGAGAGTGCAGAAAGACTTCACTAGAGGCTCTTGAGAGACAGAAAGATCTCACACCA 922
 DB 204 -----PROVALGLYLEUVALLEUGLIUPROPROTHLEUPROLEU 217

QY 923 GACAAAGCCAGATCTT-----CTTGCCCATCATCAAAACACACAGAGATCC 976
 DB 217 ISELNALALEUTHLEUARGPHELEUGLNIHISEULYTRYVALALATARGALALE- 236
 QY 977 ACCTGCCATTCGACAGCTCTTCTACATGCCAGAAAGCAGCATATGAACGAGACCA 1036
 DB 237 -----PROSETPROASPTIRALALVALHISALALEUALASERIALPHEGLYPROLEULE 255
 QY 1037 TACAGGTACATGACG-----ACG 1057
 DB 255 EULARGLIETPROSETPROSEGLYGLIUGLYASPGLYSERGLUPROVALPROASPPHEPROV 275
 QY 1058 TCCCTGCTCAGAGATGTCGCAAGCTGCAGCCAGCCAC-----AGCCCTCG 1099
 DB 275 ALLEULEULEUGLIUARGLEUVALGIN--GLIHISVALGLIUGLIUASPALALAPROPR 294
 QY 1100 AATGCTGCCCTATGAAACACAAACACTCGAGAGAACCTGACCCACACA----- 1145
 DB 294 CALALEUPRO---PROLYSPROSELYSALALALYSPROALAPROTHRALALEUALASNG 313
 QY 1146 -----AAGCTGATGAGAAAGATGCTGCGAGAAATGATGATGATGAGATACAG 1198
 DB 313 YGLYSERPROPROSETPROSELEUGLNASPALA-----GLUTRPTRYTRPGLYASPILESE 330
 QY 1199 TCCCGCAGCAGTGAAGATGTGTATTAATGAAGACAAAGATGCTATTTTGTGTCG 1258
 DB 330 TARGLIUULIUALASNGLIUARGLEU--ARGASPTIRPROASPTIRPHELEUVALAR 349
 QY 1259 AGACTGCTCTACAAATTCAGAGCAGACACATATGTTGGTGGGTATTTGGAACAA 1318
 DB 349 GASPLASERSELYSLIEGLINGLIU---TYTHRIEUTHLEUALGLYSGLYGLYAS 368
 QY 1319 GGTACATATGTAATCCCTTCTCTCGAGACCAATCAACAGATTTGGCTGGCCACAGG 1378
 DB 368 n-----ASNLYSEULEITLYSVALRPHENISARGASPTIRYGLYSPHESEGLUPR 386
 QY 1379 ACTACGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1436
 DB 386 OLEU-----THRPHECYSEVALVALGLIULEULEISERHISTYRARGHIS 401

RESULT 13
 P85A_BOVIN
 ID P85A_BOVIN STANDARD; PRT; 724 AA.
 AC P23727;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
 p85-alpha subunit) (PtdIns-3-kinase p85-alpha) (PI3K).
 GN PI3KRL.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9191567; PubMed=1707345;
 RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Larrea F., Panayotou G.,
 RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
 RA Courtneidge S.A., Parker P.J., Waterfield M.D.;
 RT Characterization of two 85 kd proteins that associate with receptor
 RT tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase.";
 RL Cell 65:91-104(1991).
 RN [2]
 RP CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.
 RX MEDLINE=93049176; PubMed=1330535;
 RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,
 RA Fry M.J., Blundell T.L., Mollmer A., Waterfield M.D.;
 RT Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
 RT domain with a PDGF receptor phosphorylation site: structural features
 RT and analysis of conformational changes.";
 RL EMBO J. 11:4261-4272(1992).

RP STRUCTURE BY NMR OF 1-84.
 RX MEDLINE=93272320; PubMed=7684655;
 RA Booker G.W., Gout I., Downing A.K., Driscoll P.C., Boyd J.,
 RT Waterfield M.D., Campbell I.D.;
 RT "Solution structure and ligand-binding site of the SH3 domain of the
 RT p85 alpha subunit of phosphatidylinositol 3'-kinase";
 RT Cell 73:813-822(1993).
 RN [4]
 RP STRUCTURE BY NMR OF 314-431.
 RX MEDLINE=92357146; PubMed=1323062;
 RA Booker G.W., Breeze A.L., Downing A.K., Panayotou G., Gout I.,
 RT Waterfield M.D., Campbell I.D.;
 RT "Structure of an SH2 domain of the p85 alpha subunit of
 RT phosphatidylinositol-3-OH kinase.";
 RT Nature 358:684-687(1992).
 RN [5]
 RP STRUCTURE BY NMR OF 321-434.
 RX MEDLINE=97110350; PubMed=8952511;
 RA Genther U.L., Liu Y., Sanford D., Bachovchin W.W., Schaffhausen B.;
 RT "NMR analysis of interactions of a phosphatidylinositol 3'-kinase SH2
 RT domain with phosphotyrosine peptides reveals interdependence of major
 RT binding sites";
 RL Biochemistry 35:15570-15581(1996).
 RL [6]
 RP STRUCTURE BY NMR OF 614-724.
 RX MEDLINE=98173872; PubMed=9512716;
 RA Siegal G., Davis B., Kristensen S.M., Sankar A., Linacre J.,
 RT Stein R.C., Panayotou G., Waterfield M.D., Driscoll P.C.;
 RT "Solution structure of the C-terminal SH2 domain of the p85 alpha
 RT regulatory subunit of phosphoinositide 3-kinase";
 RL J. Mol. Biol. 276:461-478(1998).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M61745; AAA79511.1; -;
 DR PIR, A58749; A38749.
 DR PDB, 2PNA; 31-JAN-94.
 DR PDB, 2PNB; 31-JAN-94.
 DR PDB, 2PNI; 31-OCT-93.
 DR PDB, 1PNT; 31-OCT-93.
 DR PDB, 1BFT; 25-FEB-98.
 DR PDB, 1BFJ; 25-FEB-98.
 DR InterPro: IPR001720; PI3KINase_P85.
 DR InterPro: IPR000198; RhoGAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhoGAP; 1.
 DR PRINTS; PR00678; PI3KINASEP85.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART; SM00324; RhoGAP; 1.
 DR SMART; SM00252; SH2; 2.
 DR SMART; SM00326; SH3; 1.
 DR

DR	PROSITE: PS50001; SH2; 2.	DR	
DR	PROSITE: PS50002; SH3; 1.	DR	
KW	SH3 domain; SH2 domain; Repeat; 3D-structure.	KW	
FT	DOMAIN 3 79 SH3 DOMAIN.	FT	
FT	DOMAIN 129 273 GAP DOMAIN.	FT	
FT	DOMAIN 333 428. SH2 1.	FT	
FT	DOMAIN 624 718 SH2 2.	FT	
FT	HELIIX 340 347	FT	
FT	TUNN 348 348	FT	
FT	STRAND 354 359	FT	
FT	STRAND 368 373	FT	
FT	STRAND 378 382	FT	
FT	STRAND 384 384	FT	
FT	STRAND 391 391	FT	
FT	HELIIX 401 409	FT	
FT	STRAND 413 413	FT	
FT	TUNN 418 419	FT	
FT	STRAND 422 422	FT	
FT	STRAND 427 427	FT	
SO	SEQUENCE 724 AA; 83497 MW; EBDPE6E754BBF7321 CRC64;	SO	
Alignment Scores:			
Pred. No.:	0.0536	Length:	724
Score:	137.00	Matches:	100
Percent Similarity:	37.59%	Conservative:	56
Best Local Similarity:	24.10%	Mismatches:	161
Query Match:	4.48%	Indels:	99
DB:	1	Gaps:	21
US-09-856-061-1 (1-1721) x P85A_BOVIN (1-724)			
QY	371 GTGTCAGCGGTCTC-----GGAACCACTTCGGATCAGC-----AAG	409	
DB	32 VALASNLGSLYSERLEUVALALALEUGLYPHESESAPGLYGLNGUALATLALYS	51	
QY	410 GAACCTGGCTGGGGGCCCAAGG-----GGAGAAAATGCAACGTAACAA	454	
DB	52 GLIILLEGLYTP-LEUANGLYTYRANGLUTHTHGLYGLUARGLYLSPHEPROGL	71	
QY	455 CGACTACGAGATCTCGATGTCACAGCTGCTGAGAGCATGGCCATCAATGAAATTTTACC	514	
DB	71 YHTHYTVALGIGUTYRIILEGLYARGLYSLYLSERPROBROTHRPROLSPROALGPR	91	
QY	515 AACCGAGCTATAC-----CAGGATTCGGATA	541	
DB	91 OPROALGPROLEUPROVALALAPROGLPROSERLYSTHGLUALASPERSERGLUNG	111	
QY	542 CGCAGATGACACGCTATTTCACGATATGATGAG-----GC	577	
DB	111 MLASERTHR-----LEUPROASPLEUALAGLUNGINPHEALAPROPROASPALALAPR	129	
QY	578 TCCCTTCGTGTACTCCCAAGGCTTCGTCTCC-----	611	
DB	129 OPROLEULIULEIYSLLEUVALGLUALALEGLULYSGLYLEUNGUCYSSERTHLE	149	
QY	612 -----ACTGAGACAAACACGAGGATGAGAGTGACACAGCTGGAAGAGTGACAA	664	
DB	149 UTYRATGTHRGINSERSETASINPROALGLULEARGINLEUENUSPCYSAPTH	169	
QY	665 GCCTACTCTCAAG-----GANTCGAAGCAACGCTTTAAAGATT	706	
DB	169 RLASERLEUASPLEUGLUMETPHEAPVALIHISVALLEUALASPALAPHELYSARGTY	189	
QY	707 CAATATGACAAAMATRAACACACGCTTGGCACCTCCGGCTGCTGATCACTGCC	766	
DB	189 LLEU-----LEUASPLEUFRASINPROVALLEIPEVALAL	201	
QY	767 CAGAAGTAGCAACCTTACCCCGACACCAACAGAGGAGACAGTGATCTTCGCTCC	826	
DB	201 ATALASERGLULEIILESERLEUALPROGLVALGLINSERSETGLUGUTYRIEGL	221	
QY	827 A-----AAGCCACCTTTTCAGAGATGCCAGAGGGGGCCCAAGCA	866	

Alignment Scores:

Pred. No.:	0.0911	Length:	3421
Score:	135.00	Matches:	76
Percent Similarity:	30.238	Conservative:	28
Best Local Similarity:	22.09%	Mismatches:	139
Query Match:	4.42%	Indels:	101
DB:	1	Gaps:	15

US-09-856-061-1 (1-1721) x TRGU_HSVB (1-3421)

```

QY 87 ACTCTGCTGAGGAGGAGACATGTCACACTATCTTACAGAGTGCTCCAGGATCGACC 146
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2347 ThrLeuAlaGluAspGluIleIleAlaGluLeuProTyrLeuAsnAlaAspSerLeu 2366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 147 GTGAGCCCCCTTCCAGGAGCTAGCCGTCTCAACACTAGCCCTGACTAAAGAAAGACT 206
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2367 LeuProPromeTasn-----ProAspAsp----- 2374
QY 207 GAGCAGGCTGAGCTGAGATCCCTCTCTTTTCCAGTGCCAGGACCATGACGACCCAG 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2375 -----ProIleTyr-----ThrArgValIleSerGlyThr 2384
QY 267 GGCATTAAGAGACACAGCAAGAAAGAG-----TTCGGTGATCTGAGATCCAGAACGTC 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2385 AsnIleProThrAlaThrThrGluGlySerLeuPheAlaAsp-----GlnGlnLeu 2401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 321 TCTCTGTGAAAATAGTGCATGCGCAAGCCTCAGCAGTGCCCAAGGGGGGTGCGAGCG 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2402 GluPheLeuArgProGluSerAsnProPhe----- 2411
QY 381 GTTCTGAGAACCCCTTCGGAGTCACAGAAAGAACTTGCT--GGGGTCCAGGTGAGAGA 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2412 -----ProPheAlaSerHisAspSerSerGlnSerLeuAspValProSerSerPro 2428
QY 438 AATGCAACAGTAACAGACGACTACGAGATCCT----- 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2429 SerSerGlySerAspLysTyrGluGlnAspProThrGlyIleValTyrAspAlaProVal 2448
QY 471 -----GAGTTCACCTGCTGAAGCATGCGCATCA----- 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2449 AspAspMetSerAspMetAlaMetAsnLysAlaLysAlaTrpGlnGluTrpLeuGlnAsp 2468
QY 501 -----ATGAAATTTTACCAGCCGACCTATCCAG 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2469 GlyPheAlaGluAspAspTyrArgGluLeuSerAsnAlaMetProAlaProPro---Lys 2487
QY 531 GAATCGAATACGCGATACAGCGATTTCCAGATATGAGGCGCTCCCTCTCTGTTA 590
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2488 ThrThrProValValGluSerLysGlnLysSerHisSerValAspArgAlaProThrLeu 2507
QY 591 CCTCCCAAGGCTTGTCTCCACTGAGACAAACAGGAGTGAGATGACACAGCTG 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2508 ProProLysAlaAlaProLeuProProSerAspAlaSerAlaIleMetSer----- 2524
QY 651 GAAAGAGTGACAGCCCTACCTTCAAGATGTCAAGACCAACGCTTTAAAGATCAAA 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2525 -----GlyLysProVal-----PheLys 2530
QY 711 TACACAAAATAACAGACTCTTTGGACCTCTCGGCTGCTATACACTCTCCCAAG 770
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2531 TyrThrProGlyAsnLysSerAlaValProProSerValProAla-----ProPro 2547
QY 771 AAGTACCAACCTTACCCCGACACACAGAGAGCAAGTGCATCTTGGCTCAAG 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2548 ThrLeuProProAlaProProLeuProGlnSerThrSerLysAlaAlaSerGlyProPro 2567
QY 831 CCCACCTTTCCAGAGTCCAGAGGGGGCCCAAGGACAGAGAGTGCAAAAGACTTCAGTAGG 890
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2568 ProThrLeuProProAlaProProLeuProGlnSerThrSerLys----- 2582
QY 891 GTCTTGAGACGAGAAAGATCTCACACACAGCAAGCAAGCAAGATCTTCTTGCCCATCA 950
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2583 -----AlaAlaSerGlyProProProThrLeuProProAlaProProLeuProGln 2599

```

```

QY 951 TCAAAACAAAACACAGAGAGAGTCCACCTGCCA---TTGCCAGCTCTTCTTACATGCCA 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2600 SerThrSerLysAlaAlaSerGlyProPro-ProThrLeuProProAlaProProLeuPr 2619
QY 1008 GGAAGCACCA 1017
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2619 OGInSerThr 2622

```

Search completed: April 21, 2003, 12:22:26
 Job time : 50.7582 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: Apr11 21, 2003, 12:15:31 : Search time 112.318 seconds
(without alignments)
6314.347 Million cell updates/sec

Title: US-09-856-061-1
Perfect score: 3057
Sequence: 1 acgagggccaactgcgccag.....taaaaaaaaaaaaaaaaa 1721

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 segs, 206047115 residues
Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+np.model -DEV=xlp
-Q/cgn2.1/USPTO_SPOOL/US09856061/runat_21042003_113019_3744/app_query.fasta_1.3150
-DB=SPTREMBL_21 -OEMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OJEMT=ptc -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061.eccn1.1.324.etrnat_21042003_113019_3744 -ICPU=6 -ICPU=3
-NO_XLPEXT -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMOUT=120
-NARN_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: SPREMBL_21:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organella:*
10: sp.phage:*
11: sp.plant:*
12: sp.rodent:*
13: sp.virus:*
14: sp.vertebrate:*
15: sp.unclassified:*
16: sp.virus:*
17: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2316	75.8	435	11	Q9JMJ3	Q9jMJ3 mus musculu

2	2305	75.4	435	11	Q9QZE2	Q9qze2 mus musculu
3	1132	37.0	376	4	Q9P2U9	Q9p2u9 homo sapien
4	410.5	13.4	533	11	Q922M0	Q922m0 mus musculu
5	410	13.4	525	11	Q920L0	Q920l0 ratu8 norv
6	383	12.5	530	13	Q9D607	Q9d607 gallu8 gall
7	319	10.4	456	4	Q8WV28	Q8wv28 homo sapien
8	315.5	10.3	552	13	Q9YGC1	Q9ygc1 gallu8 gall
9	309.5	10.1	433	4	Q75499	Q75499 homo sapien
10	297.5	9.7	457	11	Q9QUN3	Q9qun3 mus musculu
11	287.5	9.4	457	11	Q88504	Q88504 mus musculu
12	287.5	7.2	297	11	Q9D413	Q9d413 mus musculu
13	219.5	6.5	516	11	Q54737	Q54737 mus musculu
14	200	5.5	471	4	Q9UCX5	Q9ucx5 homo sapien
15	167.5	5.5	544	4	Q92529	Q92529 homo sapien
16	165.5	5.4	594	4	Q8TAP2	Q8tap2 homo sapien
17	165.5	5.4	594	4	Q8S466	Q8s466 y73 sarcoma
18	156.5	5.1	812	15	Q83466	Q83466 y73 sarcoma
19	154.5	5.1	217	11	Q9CX99	Q9cx99 mus musculu
20	152	5.0	486	4	Q9UCX4	Q9ucx4 homo sapien
21	151.5	5.0	2089	4	Q14676	Q14676 homo sapien
22	149.5	4.9	572	4	Q9UG31	Q9ug31 homo sapien
23	149.5	4.9	645	4	Q95928	Q95928 homo sapien
24	149.5	4.9	1254	4	Q96AT4	Q96at4 homo sapien
25	149.5	4.9	1263	4	Q9UPV8	Q9upv8 homo sapien
26	148.5	4.9	1029	12	Q9YMX0	Q9ymx0 lymantria d
27	147	4.8	2090	4	Q96QC2	Q96qc2 homo sapien
28	143.5	4.7	662	5	Q9N355	Q9n355 caenorhabd
29	142.5	4.7	1400	4	Q9UP87	Q9up87 homo sapien
30	141.5	4.6	474	11	Q61120	Q61120 mus musculu
31	141	4.6	728	4	Q9UPH9	Q9uph9 homo sapien
32	139.5	4.6	398	5	Q95PX0	Q95px0 caenorhabd
33	139.5	4.6	398	5	Q95PW9	Q95pw9 caenorhabd
34	139	4.5	17352	5	Q95YM2	Q95ym2 procambaru8
35	138.5	4.5	335	4	Q96C28	Q96c28 homo sapien
36	138.5	4.5	926	4	Q9H0K1	Q9h0k1 homo sapien
37	138	4.5	559	11	Q91252	Q91252 mus musculu
38	138	4.5	655	11	Q8R122	Q8r122 mus musculu
39	138	4.5	1385	5	Q19991	Q19991 caenorhabd
40	137.5	4.5	594	11	Q70143	Q70143 ratu8 norv
41	137	4.5	477	11	Q70142	Q70142 ratu8 norv
42	136.5	4.5	1360	16	Q988B07	Q988b07 rhlzobium l
43	135.5	4.4	498	13	Q93291	Q93291 figu rubrip
44	135	4.4	854	11	Q9QZ58	Q9qz58 mus musculu
45	134.5	4.4	309	11	Q9QV20	Q9qv20 mus sp. shb

ALIGNMENTS

RESULT 1	ID	Q9JMJ3	PRELIMINARY:	PRT:	435 AA.
AC	Q9JMJ3	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	MIST.				
GN	CLNK OR MIST.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_Taxid=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20208989; PubMed=10744659;				
RA	Goltisuka R., Kanazashi H., Sasamura H., Fujimura Y., Hidaka Y.,				
RA	Tatsuno A., Ra C., Hayashi K., Kitamura D.;				
RT	"A BASH/SLP-76-related adaptor protein MIST/Clnk involved in Ige				
RT	receptor-mediated mast cell degranulation.";				
RL	Int. Immunol. 12:573-580(2000).				
DR	EMBL; AB021220; BAA96240.1; -				
DR	HSSP; P23727; 2PNB.				
DR	MGI; MGI:1351468; Clnk.				
DR	InterPro; IPR000980; SH2.				

DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS00001; SH2; 1.
 DR SEQUENCE 435 AA; 49513 MW; 4D1BD3E2P0C61ED6 CRC64;

Alignment Scores:

Pred. No.:	1,296-197	Length:	435
Score:	2316.00	Matches:	435
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	75.76%	Indels:	0
DB:	11	Gaps:	0

US-09-856-061-1 (1-1721) x Q9JMU3 (1-435)

QY 255 ATGACACACGAGGCAATTAAGCAAGCAAGAAAGATTGCGTATCTGAGATTCCAG 314
 DB 1 MetThrSerGlnGlyAsnLysArgThrThrLysGlnGlyAspLeuArgPheGln 20
 QY 315 AACGTCTCTGCTGTAATAATAGTCATGCGCAAGCCTCAGAGAGTGGCAAGGGCGGT 374
 DB 21 AsnValSerLeuLeuLysAsnArgSerTrpProSerLeuSerSerAlaLysGlyArgCys 40
 QY 375 CGAGCGGTTCTTGGAACCACTCCGATCACAAGAAAGAACTTGCGGTGCCAGGTGA 434
 DB 41 ArgAlaValLeuGlnProLeuProAspHisArgArgAsnLeuAlaGlyValProGly 60
 QY 435 GAAATATCAACAGTAAACAGACGATCAGAGATCTGATCCAGCTCTCAAGGCTATG 494
 DB 61 GlnLysCysAsnSerAsnAsnAspTyrGlnLysProGlnLysGlnLeuLysAlaLys 80
 QY 495 CCATCAATGAAATTTTACACGACGAGACCTATCCAGAAATGGAATACGAGATACAGC 554
 DB 81 ProSerMetLysIleLeuProAlaArgProIleGlnLysGlnLysAlaAspThrArg 100
 QY 555 TATTTCCAGGATATGATGAGAGGCTCCCTTGTATACCTCCCAAGGCTTGTCTGCACT 614
 DB 101 TyrPheGlnAspMetMetGlnAlaProLeuLeuLeuProLysAlaSerValSerThr 120
 QY 615 GAGAGACAACACGAGATGTGAGATGACAGCTGGAAGAGAGTGGCAAGCCTCACTTC 674
 DB 121 GlnArgGlnThrArgAspValArgMetThrGlnLeuGlnLysValAspLysProThrPhe 140
 QY 675 AAGATGTCAGAACCAACGCTTTAAGAGATTCAATATACAAATAAACAAGACTCTCT 734
 DB 141 LysAspValArgSerGlnArgPheLysGlyPheLysTyrThrLysIleAsnLysThrPro 160
 QY 735 TTGGCACCCTCTCGGCTGCTATCCTCTCCCAAGAAATACCAAGCCTTACCCCAAGA 794
 DB 161 LeuProProAlaArgProAlaIleThrLeuProLysTyrGlnProLeuProProAla 180
 QY 795 CCACACAGAGAGAGACGATGATACCTGCTCCAAAGCCCACTTCCAGAAAGTCCAGG 854
 DB 181 ProProLysLysSerSerAlaTyrPheAlaProLysProThrPheProGlnValGlnArg 200
 QY 855 GGGCCCAAGCAGAGAGATGCAAAAGACTTCAAGAGGCTCTTGAGAGCAGAAAGAAATCT 914
 DB 201 GlyProArgGlnArgSerAlaLysAspPheSerArgValLeuGlyAlaGlnLysLysSer 220
 QY 915 CACACACAGCAAAAGCAGAAATCTTCTGCGCATCATCAACCAAAACACAGAGAGAT 974
 DB 221 HisHisGlnThrLysProGlnSerSerCysProSerSerSerSerSerSerSerSerSer 240
 QY 975 CCACCTGCGCATTCAGCTCTTCTACATGCGAGAAAGACAGTATACAAAGCAGACAGAC 1034
 DB 241 ProProAlaIleLeuAlaSerSerSerTyrMetProGlyLysHisSerIleGlnAlaArgAsp 260
 QY 1035 CATACAGATGATGATGACAGACTCTCTGCTCAGAGATGCCAAGCTCAGCCACAGCAGC 1094
 DB 261 HisThrGlySerMetGlnHisCysProAlaGlnArgCysGlnAlaAlaSerHisSer 280

QY 1095 CCTGATGCTGCTGCTATGTAACCAACCAACTCGGAGAAAGCTGACCCACAAAGCTGAT 1154
 DB 261 ProArgMetLeuProTyrGlnAsnThrAsnSerGlnLysProAspProThrLysProAsp 300
 QY 1155 GAGAAGATGCTGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1214
 DB 301 GlnLysAspValIlePheLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 320
 QY 1215 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
 DB 321 AspValLeuMetLysGlnLysAsnLysAspGlyThrPheLeuValArgAspCysSerThrLys 340
 QY 1275 TCCAAGCAGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334
 DB 341 SerLysAlaGlnProTyrValLeuValAlaPheTyrGlnLysValLysValLysValLys 360
 QY 1335 ATCCGTTCTCTGAGAGACCAATCAACAGTTGCTCCGCGACAGAGATCAAGAAATGAG 1394
 DB 361 IleArgPheLeuGlnLysSerAsnGlnGlnPheAlaLeuGlyThrGlyLeuArgGlyAsnGln 380
 QY 1395 ATGTTGATTTCTTGGAAGACATCATTTGACATACATATTTTCCATTTCTTAATA 1454
 DB 381 MetPheAspSerValGlnAspIleIleGlnHisTyrThrTyrPheProIleLeuLeuIle 400
 QY 1455 GATGGAAGACAGAGCTGCGACGAGGAAACAGTGTACCTACCCAGCCAGCTGCTCTC 1514
 DB 401 AspLysLysAspLysAlaAlaArgArgLysGlnCysTyrLeuThrGlnProLeuProLeu 420
 QY 1515 GCCAGGCTCTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTGAG 1559
 DB 421 AlaArgLeuLeuLeuThrGlnTyrSerSerGlnAlaLeuHisGln 435

RESULT 2

Q9QZE2 PRELIMINARY: PRT: 435 AA.
 AC Q9QZE2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE CLNK.
 GN CLNK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=20029811; PubMed=10562326;
 RA Cao M.Y., Davidson D., Yu J., Latour S., Veilleux A.;
 RT "Clnk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
 RT stimulated Hemopoietic Cells";
 RL J. Exp. Med. 190:1527-1534(1999).
 DR EMBL: AF187819; AAF14299.1;
 DR HSSP: P23727; 2PNB.
 DR MGD: MGI:1351468; Clnk.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS00001; SH2; 1.
 DR SEQUENCE 435 AA; 49492 MW; 5CD27EC971FC0B5 CRC64;

Alignment Scores:

Pred. No.:	1,236-196	Length:	435
Score:	2305.00	Matches:	433
Percent Similarity:	99.54%	Conservative:	0
Best Local Similarity:	99.54%	Mismatches:	2
Query Match:	75.40%	Indels:	0
DB:	11	Gaps:	0

US-09-856-061-1 (1-1721) x Q9QZE2 (1-435)

```

OY 255 ATGACCAAGCCAGGACATTAAGACAAAGAAAGAGATTGGTGATCTGACATTCCAG 314
|
|
|
Db 1 MetHsSerGlnGlnsLnysArgThrThrLysGluGlyPheGlySpleuAArgPheGln 20
OY 315 AAGGTCCTCTGCGTGAAGAAATAGTGATGGCCAAAGCTCGAGTGGCCAAAGGGCGGTG 374
|
|
|
Db 21 AsnValSerLeuLeuLysAsnArgSerTrpProSerLeuSerSerAlaLysGlyAlaGly 40
OY 375 CGAGCGCTTCGAAACCACTTCGGATCAGACAGAAAGAACTTGGCTGGGGTCCAGAGTGA 434
|
|
|
Db 41 ArgAlaValLeuGluProLeuProAspHisArgArgSnuLeuAlaGlyValProGlyGly 60
OY 435 GAAAAATGCAACAGTAAACAAGACAGTACGAAGATCTGATGTCAGCTGCTGAAGGCTGG 494
|
|
|
Db 61 GluLysCysAsnSerAsnAsnAspTyrGluAspProGluPheGlnLeuLeuLysAlaTrp 80
OY 495 CCATCAATGGAATTTACAGCCAGACCTATCCAGGAATCGGAATCGCAGATACAGC 554
|
|
|
Db 81 ProSerMetLysIleLeuProAlaArgProIleGlnLysSerGluTyrAlaAspThrArg 100
OY 555 TATTTCCAGATATATGAGAGCTCCCTCTCTGTACCTCCAGAGCTTCTGCTCCACT 614
|
|
|
Db 101 TyrPheGlnAspThrMetGluAlaProLeuLeuLeuProProLysAlaSerValSerThr 120
OY 615 GAGAGAAACCAAGGATGTGAGATGACACAGCTGGAAGAGTGGACAAAGCTTACCTTC 674
|
|
|
Db 121 GluArgGlnThrArgAspValArgMetThrHisLeuGluGluValAspLysProThrPhe 140
OY 675 AAGGATGTCAAGACCAAGCTTTAAAGATTCAATACAAAATAACAAAGACTCT 734
|
|
|
Db 141 LysAspValArgSerGlnArgPheLysGlyPheLysTyrThrLysIleAsnLysThrPro 160
OY 735 TTGCGACCTCTCGGCTGCTATGACTCTCCCAAGAAATGCCAACCTTACCCCGACA 794
|
|
|
Db 161 LeuProProArgProAlaIleThrLeuProLysLysTyrGlnProLeuProAla 180
OY 795 CCACAGAGGAGAGAGTGCATCTGCTCCCAAGCCAGCTTCCAGAGTCCAGAG 854
|
|
|
Db 181 ProProGluGlnSerSerAlaTyrPheAlaProLysProThrPheProGluValGlnArg 200
OY 855 GGGCCAGGAGAGAGAGTGCAGAAAGACTTCAAGAGGCTCTTGAAGCAGAAAGATCT 914
|
|
|
Db 201 GlyProArgGlnArgSerAlaLysAspPheSerArgValLeuGlyAlaGluGlnSer 220
OY 915 CACCAACGAGAAAGCCAGATCTTCTGCCCATCATCAAAACCAAAACACAGAGAGT 974
|
|
|
Db 221 HisHisGlnThrLysProGlnSerSerCysProSerSerAsnGlnAsnThrGlnLysSer 240
OY 975 CCACCTGCATTTGCAGCTCTTCTACATGCCAGAGAAAGCAGATATCAAGCCAGAGAC 1034
|
|
|
Db 241 ProProAlaIleIleAlaSerSerSerIleMetProGlyLysHisSerIleGlnAlaArgAsp 260
OY 1035 CATACAGGTAGCAGTGCAGCTGCTGCTGACAGATGCCAAAGTGCAGCCAGCCAGC 1094
|
|
|
Db 261 HisThrGlySerMetGlnHisCysProAlaGlnArgCysGlnAlaIleAlaSerHisSer 280
OY 1095 CCTGCAATGCTGCCCTTANGAAGAAACAAAGCTGGAGAAACCTGACCCCAAGGCTGAT 1154
|
|
|
Db 281 ProArgMetLeuProTyrGlnAsnThrAsnSerGlnLysProAspProThrLysProAsp 300
OY 1155 GAGAAAGATGCTGCAGAAATGATGATACATTTGAGAAATACAGTCCGAGGAGCTGGA 1214
|
|
|
Db 301 GluLysAspValTrpGlnAsnGlnTrpTyrIleGlyGluTyrSerArgGlnAlaValGlu 320
OY 1215 GATGCTTATATGAAAGAAACAAGATGATTTTGGTCCGAGACTGCTCTCAAAA 1274
|
|
|
Db 321 AspValIleMetLysGlnAsnLysAspGlyThrPheLeuValArgAspLysSerThrLys 340
OY 1275 TCCAAAGGAGCAACATATGTTGGTGTCTTTATGGGAACAGGCTACATGTAAGAA 1334
|
|
|
Db 341 SerLysAlaGluProTyrValLeuValAlaPheTyrGlyAsnLysValTyrAsnValLys 360

```

```

OY 1335 ATCCGTTTCCTGAGAGACATCAACAGTTTGCCCTGGGACAGAGACTACAGGAATGAG 1394
|
|
|
Db 361 IleArgPheLeuGlnSerAsnGlnGlnPheAlaLeuGlyThrGlyLeuArgGlyAsnGlu 380
OY 1395 ATGTTTCATTTCTGTGAGACAGATCATTTGAACACTACATATTTTCCATTCTGTAATA 1454
|
|
|
Db 381 MetPheAspSerValGlnLysPheIleIleGlnHisTyrThrTyrPheProIleLeuLeuIle 400
OY 1455 GATGGAAAGACAAAGCTGACGACGACAGAAACAGTGTACTCTACCCAGCCAGCTCTTC 1514
|
|
|
Db 401 AspGlyLysAspLysAlaAlaArgArgLysGlnCysTyrLeuThrGlnProLeuProLeu 420
OY 1515 GCCAGGCTCCTCTCACTGACTGACTCCAGCAGCAGCACTTCATGAG 1559
|
|
|
Db 421 AlaArgLeuLeuLeuThrGlnTyrSerSerGlnAlaLeuHisGlu 435

```

RESULT 3

ID	Q9P2U9	PRELIMINARY:	PRT:	376 AA.
AC	Q9P2U9			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	MIST (Fragment).			
GN	MIST.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20208989; PubMed=10744659;			
RA	Goltsuka R., Kanazashi H., Sasamura H., Fujimura Y., Hidaka Y.,			
RA	Tatsuno A., Ra C., Hayashi K., Kitamura D.;			
RT	"A BSH/SLP-76-related adaptor protein MIST/Clk involved in Ige			
RT	receptor-mediated mast cell degranulation.";			
RL	Int. Immunol. 12:573-580(2000).			
DR	EMBL; AB032369; BAA96241.1; -			
DR	HSP; P23727; ZPNB.			
DR	InterPro: IPR000980; SH2.			
DR	Pfam: PF00017; SH2; 1.			
DR	PRINTS; PR00401; SH2DOMAIN.			
DR	PRODOM; PD000093; SH2; 1.			
DR	SMART; SM00252; SH2; 1.			
DR	PROSITE; PS50001; SH2; 1.			
FT	NON_TER 1 1			
FT	NON_TER 376 376			
SO	SEQUENCE 376 AA; 43542 MW; 389421B629B02B2 CRC64;			

Alignment Scores:

Pred. NO.:	4.05e-92	Length:	376
Score:	1132.00	Matches:	229
Percent Similarity:	72.07%	Conservative:	42
Best Local Similarity:	60.90%	Mismatches:	101
Query Match:	37.03%	Indels:	4
DB:	4	Gaps:	3

US-09-856-061-1 (1-1721) x Q9P2U9 (1-376)

```

OY 309 TTCGAGACGCTCTCTCTGCTGAAAATAGTCATGAGCCAGCTCAGCAGTGGCCAAAGG 368
|
|
|
Db 1 PheGlnAsnPheSerLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGly 20
OY 369 CGGTGTCGAGGCTTCTGGAACCACTTCGATCAGACAGAACTTGGCTGGGGTCCCA 428
|
|
|
Db 21 GlnTyrGlnArgMetAsnLysProLeuLeuAspTrpGluArgAsnPheAlaValLeu 40
OY 429 GGTGAGAAAATGCAACAGTACAGCACTACAGATCTCTGAGTTCACAGTCTGAGAG 488
|
|
|
Db 41 AspGlyAlaLysGlnHisSerAspAspArgTyrAspAspProGluLeuArgMetGlu 60
OY 489 GCATGGCCATCAATGAAATTTTACCAGCAGACTTACAGGAATCGGAATACAGCAT 548
|
|
|

```

```

Db 61 ThTrpInSerIleLysIleLeuProAlaArgProIleLysIleuSerGluTyrAlaAsp 80
QY 549 ACACGCTATTTCAGGATATGATGAGGCGCTCTGTACTCCCAAGCTTCTGTC 608
Db 81 ThH1eTyrPheLysValAlaMetAspThrProLeuProLeuAspThrArgTrpSerIle 100
QY 609 TCCACTGAGACAAACAGGAGTGTAGATGACACACTGAGAGATGACAGACCT 668
Db 101 SerIleGlyGlnProThrTrpAsnThrGln---ThArgLeuGlnArgValAspLysPro 119
QY 669 ACCTTGAAGATGTCAGAACCCCAAGCTTTAAAGATTCAATAACAAATAAACAAG 728
Db 120 IleSerArgPylValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLys 139
QY 729 ACTCTTTTCCACCTCTGCGCTGTATCACTCTCCCAAGAGTACCAACCTTACCC 788
Db 140 IleProLeuProProArgProArgProLeuIleThrLeuProLysLysTyrGlnProLeuPro 159
QY 789 CCAGCACCCAGAGAGAGACAGTGCATCTTCCCTCAAGCCCACTTTCCAGAACTC 848
Db 160 Pro---GluProGluSerSerArgProProLeuSerGlnArgHisThrPheProGluVal 178
QY 849 CAGAGGGGGCCAGGAGAGAGAGTGCAAAAGACTTCAGTAGGGTCTTGAGACAGAGA 908
Db 179 GlnGlyMetProSerGlnIleSerLeuArgAspLeuSerGlnValIleuGlnAlaGluLys 198
QY 909 GAATCTCACACACAGCAACAAAGCAAGATCTTCTGCGCCATCAACCAACCAACAGC 968
Db 199 ValProHisAsnGlnArgLysProGluSerThrHisLeuGlnAlaAsnGlnAsnThrGln 218
QY 969 AAGAGTCCACCTGCCATTCGCCAGCTCTTCTCATCATGCCAGAAAGCATATACAGCC 1028
Db 219 GluIleProLeuAlaIleSerSerSerPheThrThrSerAsnHisSerValGlnAsn 238
QY 1029 AGAGACATATACAGATGAGATGAGACATGCTCTGCTCAGAGATGCCAAGCTGCACCAAC 1088
Db 239 ArgAspHisArgGlyGlyMetGlnProCysSerProGlnArgCysGlnProProAlaSer 258
QY 1089 CACAGCCCT---CGAATGCTGCCCTATGAAAACACAACTCCGAAACCTGACCC 1142
Db 259 CysSerProHisGlnAlaAsnIleLeuProTyrLysTyrThrSerTrpArgProProPhePro 278
QY 1143 ACAAGCCGTGATGAAAGATGCTGCGCAAGATGATGATGATGGAATATACAGTCCG 1202
Db 279 LysArgSerAspArgLysAspValGlnHisAsnGluTyrGlyIleGlyLysSerArg 298
QY 1203 CAGGCACTGAGAGATGCTGTTATGAAAGAGAAAGAGATGATGATGTTTGTGCGAGAC 1262
Db 299 GlnAlaValGlnGluAlaPheMetLysGlnAsnLysAspGlySerPheLeuValArgAsp 318
QY 1263 TGCCTTACAAATCCAGGCAAGACCATATGTTTGTGTTGTTTATGGAACAGGCTC 1322
Db 319 CysSerThrLysSerLysGlnGluProTyrValLeuAlaValPheTyrGlnAlaAsnVal 338
QY 1323 TACATGTGAATAATCCGTTCTCTGAGAGCATCAACAGTTTGGCCCTGGGACAGACTA 1382
Db 339 TyrAsnValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlyThrGlyLeu 358
QY 1383 CGAGGAATGAGATGTTGATCTGTGAGAGACATCATTAACACTAC 1430
Db 359 ArgGlyAspGlnLysPheAspSerValGlnLysPheIleGlnHisTyr 374

```

RESULT 4

```

0922M0 PRELIMINARY: PRT: 533 AA.
AC 0922M0:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
CN Lymphocyte cytosolic protein 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC006948; AA06948.1; -.
DR MGD: MGI:1321402; Lcp2.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR PROSITE: PS0001; SH2; 1.
SQ SEQUENCE 533 AA; 60238 MW; 50AEAD025EF0AD01 CRC64;

Alignment Scores:
Pred. No.: 8,896-28 Length: 533
Score: 410.50 Matches: 148
Percent Similarity: 39.68% Conserved: 50
Best Local Similarity: 29.66% Mismatches: 176
Query Match: 13.43% Indels: 125
Db: Gaps: 20

US-09-856-061-1 (1-1721) x 0922M0 (1-533)
QY 294 TTCGGTATCTGAGATTCACAGACTCTCTGCTG-----AAAATAGC 338
Db 57 PheProLysLeuArgMetProLeuLeuSerLysLeuSerGlnAsnLysAsnGln 76
QY 339 TCATGGCCAGGCTCAGCAGCAGCAAGGCGGTGTCAGGCGGTTCGGAACCACTTCG 398
Db 77 GluArgArgSerIlePheThrArgLysProGlnIleProAlaPheLeuGlnGluThrGlu 96
QY 399 GATCACAGAGAACTTGCTGGCG-----422
Db 97 SerHisGlnGlnAspAspGlyGlyTyrSerSerPheGlnAspArgTyrGluSerProAsn 116
QY 423 -----GTCCAGGTGGGAAATAATGCACAGTAACACGACTCGAAGATCT- 470
Db 117 AspAspAspProAspGlyGlu-----AspAspGlyAspTyrGluSerProAsnGln 133
QY 471 GAGTTCACGCTGCTG-----AAGCATGGCCATCA-----500
Db 134 GluGlnGlnAlaLeuValAspAspAlaAlaAspTyrGluProProSerAsnAsnGln 153
QY 501 -----ATGAATAATTTACCAAGCCAGCAGCTATCCAG-----GAATCGGAATAC 542
Db 154 GluAlaLeuGlnSerSerIleLeuProProAsnSerPheHisAsnThrAsnSerMetTyr 173
QY 543 GCAGATACACCCCTATTTCAGAGATATGATGAGAGGCTCCCTTCTGTAACT- 596
Db 174 IleAspArgProProThrGlyLysValSerGlnGlnProProValProProLeuArgPro 193
QY 597 AAGCTCTCTGTC-----608
Db 194 LysProAlaLeuProProLeuProThrGlyArgAsnHisSerProLeuSerProHis 213
QY 608 -----608
Db 214 ProAsnHisGlnGluProSerArgSerGlyAsnAsnLysThrAlaLysLeuProAlaPro 233
QY 609 TCCACTGAGAGACAAACAGGAGATGAGAGATGACACAGCTGGAAGAGTGAC-----662
Db 234 SerIleAspArgSerThrLysProProLeuAspArgSerLeuAlaProLeuAspArgGlu 253
QY 663 -----AAGCTTACTTCAAGATGTCGAAGCCCAACCTTTAAAGA 704
Db 254 ProPheIleuGlnLysLysProProPheSerAspLysProSerAlaProLeuGlyArg 273
QY 705 TTCAAATACAAATAAACAAGACTCTTCCACACTCTGCGCTGATACACTCTC 764
Db 274 GluHisLeuProLysIleGlnLysProProLeuPro-----ProAlaMetAspArg 290
QY 765 CCCAAGAAGTACCAACCTTAACCCCGACAGCACCAAGAGAGAGAGCATGCTTCGCT 824

```



```

Db 291 HisGluArgAsnGluValGluGlyProValThrThrArg----- 303
      :::::  ::  |||  |||
Qy 825 CCAAGCCACACCTTCCAGAGTCCAGAGGGGCCAGAGAGTGCAGAAAGACTTC 884
      |||||  |||  |||||  |||
Db 304 ---LysProProValProArgHisGlyArgGlyProAspArgGlu----- 318
      |||||  |||  |||||  |||
Qy 885 AGTAGGTCCTTGAGCAGGAAGAGATTCACACCGACAGAAAGCCAGATCTTCTGC 944
      |||||  |||  |||||  |||
Db 319 -----AsnAspGluAspAspValHisGlnArgProLeuProGlnProSerLeu 334
      |||||  |||  |||||  |||
Qy 945 CCATCATCAACCAACACACAGAGAGTCCACCTGCCATTCGACAGCTC----- 995
      |||||  |||  |||||  |||
Db 335 ProSerMetSerSerAsnThrPheProSerArgSerValGlnProSerSerLysAsnThr 354
      |||||  |||  |||||  |||
Qy 996 -----TCCTACATGCCAGAGAAAGCAGATATACAGCCAGAGACCATACAGGTAGC 1046
      |||||  |||  |||||  |||
Db 355 PheProLeuAlaHisMetProGlyAlaPheSer-----GluSerAsnIleGlyPhe 371
      |||||  |||  |||||  |||
Qy 1047 ATGCAGCACTGT-----CTGCTCAGAGATGCCAGCTGCAGCCAGCCAGACGCTC--- 1097
      |||  |||  |||  |||
Db 372 GlnGlnSerAlaSerLeuProProTyrPheSerGlnGlyProGlyAsnArgProProLeu 391
      |||  |||  |||  |||
Qy 1098 -----CGAATGCTGCCCTATGAGAAACACAACTCGGAGAACTGCACCCACCA 1145
      |||  |||  |||  |||
Db 392 ArgSerGlnGlyArgAsnLeuProLeuProValProAsn-----ArgProGlnProPro 409
      |||  |||  |||  |||
Qy 1146 AAGCTGTATGAGAGAGATGTC---TGCAGAGATGATGTCATGTCAGATACAGTGC 1202
      |||  |||  |||  |||
Db 410 SerProGlnGlnGluGlnThrProLeuAspGlnIleTyrPheValSerTyrIleThrArg 429
      |||  |||  |||  |||
Qy 1203 CAGGACAGTGCAGATGTGTATGTAAGAGACAGAGATGACTTTTGTGTCGAGAC 1262
      |||  |||  |||  |||
Db 430 ProGlnAlaGlnAlaAlaLeuArgLysIleAsnGlnAspGlyThrPheLeuValArgAsp 449
      |||  |||  |||  |||
Qy 1263 TGCCTTACAAATCCAGGACAGACATATGTTTGTGTCGTTTGTGTCGAGAC 1322
      |||  |||  |||  |||
Db 450 SerSerLysThrAlaAsnAsnProTyrValLeuMetValLeuTyrLysAspLysVal 469
      |||  |||  |||  |||
Qy 1323 TACATGTGAATCCGCTTCCGAGAGCATACAACTTCCCTGCGGACAGACATA 1382
      |||||  |||  |||||  |||
Db 470 TyrAsnIleGlnIleArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 489
      |||||  |||  |||||  |||
Qy 1383 CAGAGAAATGATGTTGATTCGTGAGAGACATCATGTAACACTACATATTTTCCC 1442
      |||||  |||  |||||  |||
Db 490 ArgGlyLysGlnAspPheLeuSerValSerAspIleIleAspTyrPheArgLysMetPro 509
      |||||  |||  |||||  |||
Qy 1443 ATTCTGCTAATAGATGGAGAAAGCAAGCTGCACGACGAGAAACAGTCTACCTCACC 1499
      |||||  |||  |||||  |||
Db 510 LeuLeuLeuIleAspGlyLysAsnArgLysSer---ArgTyrGlnGlySerThrLeuThr 527
      |||||  |||  |||||  |||

```

```

DR PROSITE: P550001; SH2; 1.
SQ SEQUENCE 525 AA; 59521 MW; 0B1CDD40DABE7372 CRC64;

Alignment Scores:
Pred. No.: 9.81e-28 Length: 525
Score: 410.00 Matches: 127
Percent Similarity: 44.72% Conservative: 51
Best Local Similarity: 31.91% Mismatches: 148
Query Match: 13.41% Indels: 72
DB: 11 Gaps: 16

US-09-856-061-1 (1-1721) x Q920L0 (1-525)
Qy 456 GACTATGAGAGATCCGTGAGATTCACCTGCTGAGAGCATGCGCATATGAAATTTTACCA 515
      |||||  |||  |||||  |||
Db 144 AspTyrGlnProProProProSerAsnAspGlnGlnAlaLeuGlnAsn---SerIleLeuPro 162
      |||||  |||  |||||  |||
Qy 516 GCCAGACCTATC-----CAGGAATCGGAATACGACATACAGCCTATTTCCAGGATATG 569
      |||||  |||  |||||  |||
Db 163 AlaLysProPheProAsnThrAsnSerMetTyrIleAspArgProProThrGlyLysVal 182
      |||||  |||  |||||  |||
Qy 570 ATGAGAGCTCCC-----CTTCTGTACCTGCC----- 596
      |||  |||  |||  |||
Db 183 SerGlnGlnProProValProProGlnArgProMetAlaAlaLeuProProLeuProThr 202
      |||  |||  |||  |||
Qy 597 -----AAGCTTCTGCTCCTCAGTGCAGAGCAAAACGAG 629
      |||  |||  |||  |||
Db 203 GlyArgAsnHisSerProAsnHisGlnGlnLysProSerArgAsnAlaGlnHisLysThrAla 222
      |||  |||  |||  |||
Qy 630 GATGTAGAGATGACACACCTGGAAGAGTGCAGAAAGCTTCTTCAAGATGTACAGAAC 689
      |||  |||  |||  |||
Db 223 LysLeuProAlaProSerIleAspArgSerThrLysProProLeuAsp-----ArgSer 240
      |||  |||  |||  |||
Qy 690 CAAGCTTTAAAGATTCAAATACACAAATTAACAGAGCTCTTGTG----- 737
      |||  |||  |||  |||
Db 241 LeuAla-----ProProLeuAspArgGlnProLeuLeuGlnLysLys 254
      |||  |||  |||  |||
Qy 738 CCACCTCTCGGCTGTATCATCTCTCCCAAGAGTACCAACCC-----TTACCC 788
      |||  |||  |||  |||
Db 255 ProSerAspLysProSerAlaLeuLeuGlnArgGlnHisLeuProLysIleGlnLysPro 274
      |||  |||  |||  |||
Qy 789 CCAGACACACCA-----GAGAGAGCAGAGTCATCTGCT 824
      |||  |||  |||  |||
Db 275 ProLysProProAlaMetAspArgHisGlnArgAsnGlnArgIleGlyProLeuThrAla 294
      |||  |||  |||  |||
Qy 825 CCAAGGCCACCTTCCAGAAAGTCCAGAGGGGCCAGAGAGAGTGCAGAAAGACTTC 884
      |||||  |||  |||||  |||
Db 295 ArgLysProProValProArgHisGlyArgGlyProAspArgArgAsp----- 310
      |||||  |||  |||||  |||
Qy 885 AGTAGGTCCTTGAGCAGAGAGATTCACACCGACAGCAAGCCAGATCTTCTGC 944
      |||||  |||  |||||  |||
Db 311 -----AsnAspLysAspAspValHisGlnArgProLeuProHisProSerLeu 326
      |||||  |||  |||||  |||
Qy 945 CCATCATCAACCAAC-----ACACAGAGAGTCCACCTGCCATT 986
      |||||  |||  |||||  |||
Db 327 ProSerMetSerSerAsnThrPheProSerArgSerTyrThrLysProSerTyrAsnThr 346
      |||||  |||  |||||  |||
Qy 987 GCCAGCTTCTTACATGCGCAGAGAAAGCAGATATACAGCCAGAGACCATACAGGTAGC 1046
      |||||  |||  |||||  |||
Db 347 PheProLeuProHisMetProGlyAlaIleSer-----GluSerAsnIleGlyPhe 363
      |||||  |||  |||||  |||
Qy 1047 ATGCAGCACTGT-----CTGCTCAGAGATGCCAGTGCAGCCAGCCAGCCTC--- 1097
      |||||  |||  |||||  |||
Db 364 GlnGlnSerAlaSerLeuProSerTyrPheSerGlnGlyProSerSerArgProProVal 383
      |||||  |||  |||||  |||
Qy 1098 -----CGAATGCTGCCCTATGAGAAACCAAACTCGGAGAAACCTGACCCACCA 1145
      |||||  |||  |||||  |||
Db 384 ArgAsnGlnValArgAsnLeuProLeuProValProAsnArgProGlnPro---ProSer 402
      |||||  |||  |||||  |||
Qy 1146 AAGCTGTAGAGAGATGTCGCGAGATGATGTCATGTCAGATACAGATGCGCCAG 1205
      |||||  |||  |||||  |||
Db 403 ProGlnGlnGlnGlnSerProLeuAspGlnIleTyrPheValSerTyrIleThrArgPro 422
      |||||  |||  |||||  |||

```

D	b	160	ephePrOialaLysSerLeuAlaLsmnSnrhAsprYrIlleAspargPrOProthrrSerAr	160
			:	
Q	y	359	TGCCAAAGGGCGGTGTGCAGCGGTTCTGGACCCTTCGGATGACAGACGAACCTTGCC	418
			: :	
D	b	180	gSerSerhISgLn-----ProPrOValPrOProGlnArgPrOgLyProse	195
			: :	
Q	y	419	TGGGGTCCCA-----GGTGGAGAAAATATGCAACGTAACACAGCATACGAAGATCC	469
			:	
D	b	195	rPrOialPrOialSerPhegLYeLYArgGLYAlaSer-----	207
			:	
Q	y	470	TGACCTTCACCTGCTGAGACGATGCCATCATATGAATTTTTCACACGACCTATCC	529
			: :	
D	b	208	-----LeuPrOialPhePrOProleuPrOgLYlsmnSnrhAsprgSnnLYlly	223
			: :	
Q	y	530	GGAATGGAAATACGCGATACACGCTATTTCAGATATGATGAGAGCTCCCTCTGT	589
			: :	
D	b	223	sPrOserLYs-----	226
			:	
Q	y	590	ACCTCCCAAGGCTTCTGTCTCCACTGAGAGACAAACCGGATGTAGATGACACAGCT	649
			: :	
D	b	227	-PrOPrOialPrOserIle-----AspargSerThLYs-----ProPrOle	240
			: :	
Q	y	650	GGAAGCAAGTGGACAAACCTACCTTCAGAGATGTCAAGACCAACGCTTAAAGATTCAA	709
			:	
D	b	240	uAspargLeuGLYrPrO-----	246
Q	y	710	ATACACAAATATTAACAAAGACTCCT-----TGCC	739
			: :	
D	b	247	-PhegLYuArgLYuAsnPrOValPrOgLYATgLYsPrOgLYhISPrOgLYlYSLeuLnh	266
			:	
Q	y	740	ACCTCCTGGCGCTCTATCTACTTCCCAAGAAATACACCTTACCCCAACACC	799
			:	
D	b	266	rPrOgLYuAsnArgLYaLeuGLYgLYgLYuLnuAlaMeLeuPrOgLYsPrOPrOValPrObr	286
			:	
Q	y	800	AGAGAGACGACGATGATCTTGGCTCCAAAGCCCACTTTCAGAA-----GTCCAGAG	853
			:	
D	b	286	OsErAspargLYrGLYuArgLYAsnPrOserPrOleuArgLYgLnIlePrOValLYeLI	306
			:	
Q	y	854	GGGCCCCAGGCGAGAGGATGCAAAAGACTTCAGTAGGGTCTTGAGCAGAAAGAAATC	913
			: :	
D	b	306	nGLYrPrAlaGLYnGLYAsnArgPrOgLY-----GLYuGLYuAs	319
			:	
Q	y	914	TCAC-----CACGACGAACAAAGCAGAAATCTTCTGCCATATCATCAACCAACACA--	965
			: :	
D	b	319	phtISrIlePrOgLYnArgLYaLPrOgLYlIleSerLeuPrOProTYrSerSerAsnThrPh	339
			: :	
Q	y	966	-----CAGAGCTCCACCTGCCATATGCGACGCTCTCTATCATCCAGGAAA	1012
			: :	
D	b	339	ePrOserLYsErIleLYsAlaPrOPrOgLYsErAsnSer--IlePrOgLYAl	358
			: :	
Q	y	1013	GCAGAGTATACAGCCAGACAGACCATACAGTACATGCACACTGCTCTCAGAGATG	1072
			: :	
D	b	358	aGLYuAsnArgLYrGLYrGLYAsnSerIleSerIleSerIleu-----PrOPrOArgPhePr	375
			: :	
Q	y	1073	CCAACTGCACACCGACCAAGCCCTCGAATG-----CTGCC	1108
D	b	375	OLEuLYlsmnSnrhErArgSerPrOserArGLYrPrAlaAspLeuArgPrOProleuPr	395
Q	y	1109	CTATGAAAACACAAACTGGAGAAACCTGACCCACAAAGCCTGATGAGAAAGATCTTG	1168
			: :	
D	b	395	OllePrOserArGLYnThrAlaISnGLYnThrAsnThrGLYuAspGLY--AspSerIe	414
			: :	
Q	y	1169	GCAGAGTAAAGGATACGATGAGATATGACCTGCCACGAGCTGGAGAAAGATGTATTATGA	1228
			: :	
D	b	414	uAsnSPrOILrPrYrYAlaLYrILleSerArPrOgLYuAlaGLYuAlaAlaLeuArgLY	434
			: :	
Q	y	1229	AGAGAACAGAGATGTACTTTTTGGTCGAGACTGCTTACAAATGCAAGCGACGAAC	1288
			: :	
D	b	434	sILEuSnyAspLYrPhrPheLeuValArGAspSerSerArGLYsThrThrThrThrISPr	454
			: :	
Q	y	1289	ATATGTTTTGTGTGTTTTATGGAACAAGCTTACAAATGTGAATAATCCGTTTTCTGCA	1348
			: : :	

RT Interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
 RT receptor activation.";
 RL J. Biol. Chem. 272:27362-27368(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98361304; PubMed=9697839;
 RA Fu C., Turk C.W., Kurosaki T., Chan A.C.;
 RT "BLNK: a central linker protein in B cell activation.";
 RL Immunity 9:93-103(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20050956; PubMed=10583958;
 RA Minggishl Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
 RA Campana D., Chan A.C., Conley M.E.;
 RT "An essential role for BLNK in human B cell development.";
 RL Science 286:1954-1957(1999).
 DR EMBL: AF068180; AAC39636.1;
 DR EMBL: AF180756; AAF20382.1;
 DR EMBL: AF180740; AAF20382.1; JOINED.
 DR EMBL: AF180741; AAF20382.1; JOINED.
 DR EMBL: AF180742; AAF20382.1; JOINED.
 DR EMBL: AF180743; AAF20382.1; JOINED.
 DR EMBL: AF180744; AAF20382.1; JOINED.
 DR EMBL: AF180745; AAF20382.1; JOINED.
 DR EMBL: AF180746; AAF20382.1; JOINED.
 DR EMBL: AF180747; AAF20382.1; JOINED.
 DR EMBL: AF180748; AAF20382.1; JOINED.
 DR EMBL: AF180749; AAF20382.1; JOINED.
 DR EMBL: AF180750; AAF20382.1; JOINED.
 DR EMBL: AF180751; AAF20382.1; JOINED.
 DR EMBL: AF180752; AAF20382.1; JOINED.
 DR EMBL: AF180753; AAF20382.1; JOINED.
 DR EMBL: AF180754; AAF20382.1; JOINED.
 DR EMBL: AF180755; AAF20382.1; JOINED.
 DR HSSP: P29354; 1BM8.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 SO SEQUENCE 456 AA; 50466 MW; 95FID5485D03D397 CRC64;

Alignment scores:
 Pred. No.: 1 21e-19 Length: 456
 Score: 319.00 Matches: 99
 Percent Similarity: 39.65% Conservative: 58
 Best Local Similarity: 25.00% Mismatch: 139
 Query Match: 10.44% Indels: 100
 DB: 4 Gaps: 12

US-09-856-061-1 (1-1721) x 075498 (1-456)

QY 444 AACAGTACACGACTACGAGATCTGAGTTCAGCTGCTGAGAGCATGGCCATCATG 503
 DB 91 AsnAlaAspAspSerTYRGLuProProValGlu-----GlnGluThrArg 106
 QY 504 AAAATTTTACAGCCAGCATTCAGGATCGAATACGACAGTACACGCTATTTCCAG 563
 DB 107 ProValHisProAlaLeuProPheAlaArgGlyGluTyrlleAspAsnArgSerSerGln 126
 QY 564 GATATGATGAGCTGCC-----CTTCTGTACTCTCCCAAGGCTTGTCTTCACAT 614
 DB 127 -----ArgHisSerProProPheSerLysThrLeuProSerLysProSerTyrProSer 144
 QY 615 GAGAGACAAACAGGAGTGTGAGATGACACAGCTGGAGAGTGGACAGCTTACCTTC 674
 DB 145 GluTysAlaArgLeuThrSerThrLeuProAlaLeuThrAlaLeuGlnLysProGlnVal 164
 QY 675 AAGGATTCAGAAAGCCACGCTTTAAAGATTCAATATACAAATAAACAAGACTCCT 734
 DB 164 ----- 164
 QY 735 TTGCACACTCTCGGCTGTCTATCACTCTCCCAAGAGATACCAACCTTACCCCGACGA 794

DB 155 -----ProProLysProLysGlyLeuLeuGlnAspGlnAlaAspTyrValValProVal 182
 QY 795 CCACCAGAGAGAGAGTGCATACCTTCGCTCCAAAGCCACCTTCCAGAAAGTCCAGAG 854
 DB 183 GluAspAsnAspGluAsnTyrIleHisProThrGluSerSerProProGluLys 202
 QY 855 GGGCCCGAGGAGAGAGTGCAAAGACATTCAGTAGGCTCTTGAGCAGAAAGAAATCT 914
 DB 203 AlaPromet-----Val 206
 QY 915 CACCCACGACAAAGCCAGATCTTCTGCCATCATCAACCAAGACACA----- 965
 DB 207 AsnArgSerThrLysProAsnSerSerThrProGluSerProGlyThrAlaSerGly 226
 QY 966 -----CAGNAGAGTCCACCTGCTGATGCGAGCTCTTCTACATG 1004
 DB 227 ArgAsnSerGlyAlaTrpGluThrLysSerProPro-----AlaAlaProSerProLeu 245
 QY 1005 CCA-----GGAAGCAC-----AGTATACAAAGCAGAGAC 1034
 DB 246 ProArgAlaGlyLysLysProThrThrProLeuLysThrThrProValAlaSerGln 265
 QY 1035 CATACAGGTACCATG-----CAGCATGTCTGCTGAGATGCCAAGCTGCA--- 1082
 DB 266 AsnAlaSerSerValCysGluGluLysProLeuProAlaGluAlaGlnHisArgLysSer 285
 QY 1083 -----GCCAGCCACAGCCCTCGAATGCTGCCATATGAAACACAACTCCGAG 1130
 DB 286 HisArgGlnGluAlaValGlnSerProValPheProProAlaGlnLysGlnIleHisGln 305
 QY 1131 AAACCTGACCCC----- 1142
 DB 306 LysProLeuProLeuProAlaArgPheThrGluGlyLysAsnProThrValAspGlyProLeu 325
 QY 1143 -----ACAAAGCCTGATGAGAAAGATCTCTGACAGATGAA 1178
 DB 326 ProSerPheSerSerAsnSerThrIleSerGluGlnGluAlaGlyValIleCysLysPro 345
 QY 1179 TGTACATTTGAGAGATACAGTCCGCCAGCAGTGGACATGTGTTAATGAAGACACAG 1238
 DB 346 TrpTyrAlaGlyAlaCysAspArgLysSerAlaGluAlaGlnAlaLeuHisArgSerAsnLys 365
 QY 1239 GATGCTACTTTTGTGTCGAGACTGCTACAAATAATCCAAAGCAGAACCATATGTTTG 1298
 DB 366 AspLysSerPheLeuIleArgLysSerSerGlyHisAspSerLysGlnProTyrThrLeu 385
 QY 1299 GTGCTGTTTATGGAAACAAAGTCTACATGTGAAATCCGTTTCTCGAGACCAATCAA 1358
 DB 386 ValValPhePheAsnLysArgValTyrAsnIleProValArgPheIleGluAlaThrLys 405
 QY 1359 CAGTTGCTCCCTGGGCGACAGACTACGAGAAATGAGATGTTTATCTCTGGAAGCATC 1418
 DB 406 GlnTyrAlaLeuGlnArgLysLysAsnGlyGluGluTyrPheGlySerValAlaGluIle 425
 QY 1419 ATGAAACACTACATATTTTCCATCTGCTATATAGATGGAAAGAC 1466
 DB 426 IleArgHisGlnHisSerProLeuValLeuIleAspSerGlnAsn 441

RESULT 9
 ID 09YGC1 PRELIMINARY; PRT; 552 AA.
 AC 09YGC1;
 DT 01-MAY-1999 (TREMBLER, 10, Created)
 DT 01-MAY-1999 (TREMBLER, 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLER, 20, Last annotation update)
 DE B cell linker protein BLNK.
 GN BLNK OR BASH.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NX NCBI_TaxID=9031;

```

RN      [1]
RX      SEQUENCE FROM N.A.
RM      MEDLINE=9916381; PubMed=10023776;
RA      Ishii M., Kurosaki M., Papay R., Okawa K., Ronko I., Fu C.,
RT      Shibata M., Yamatsu A., Chan A.C., Kurosaki T.;
RL      "PLNK required for coupling syk to PLC gamma 2 and Rac1-JNK in B
        cells.";
        Immunity 10:117-125(1999).
        [2]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=99049801; PubMed=9834055;
RA      Gelsitsuka R., Fujimura Y., Kamada H., Umeda A., Morimura T.,
RA      Uetsuka K., Doi K., Tsuji S., Kitamura D.;
RT      "BASH, a novel signaling molecule preferentially expressed in B cells
        of the bursa of Fabricius.";
RL      J. Immunol. 161:5804-5808(1998).
DR      EMBL; AF089727; AACD2783.1; -.
DR      EMBL; AB015289; BAA36275.1; -.
DR      HSSP; P08487; 2PLD.
DR      InterPro: IPRO02965; P_rich_extensn.
DR      InterPro: IPRO00980; SH2.
DR      Pfam; PF00017; SH2_1.
DR      PRINTS; PRO1217; PRICHEXTENS.
DR      PRINTS; PRO0401; SH2DOMAIN.
DR      ProDom; PD000093; SH2; 1.
DR      SMART; SM00252; SH2; 1.
DR      PROSITE; PSS0001; SH2; 1.
DR      PROSITE; PSS0001; SH2; 1.
SQ      SEQUENCE 552 AA; 61822 MW; FB32179BE38D072 CRC64;

```


DT 01-MAR-2000 (TREMBLeref. 13, last sequence update)
DT 01-MAR-2002 (TREMBLeref. 20, last annotation update)
DE SH2-containing leukocyte protein 65 (lymphocyte antigen 57).
GN SLP-65 OR LY57.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN-BALB/C; TISSUE-LYMPHOID;
RC MEDLINE=98327271; PubMed=9705962;
RA Wiensands J., Schwikert J., Wollisch B., Juma H., Nielsen P.J.,
RA Beth M.;
RT "SLP-65: A new signalling component in B lymphocytes which requires
RT expression of the antigen receptor for phosphorylation.";
RL J. Exp. Med. 188:791-795(1998).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN-BALB/C; TISSUE-LYMPHOID;
RA Wiensands J., Larbolette O., Beth M.;
RT "Evidence for a pleiotropic transducer complex organized by the B cell
RT antigen receptor";
RL Proc. Natl. Acad. Sci. U.S.A. 93:7865-7870(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Okamoto N., Hayashi K., Tsuji S., Goitsuka R., Kitamura D.;
RT "BASH: B lymphocyte adaptor protein containing SH2 domain.";
RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RP TRANSPOSON-ETN;
RC Nielsen P.J., Guenet J.L.;
RA "The murine SLP-65 gene.";
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB015290; BAA34944.1; -;
DR EMBL; AJ298054; CAC1865.1; -;
DR HSSP; P23727; 1BFI.
DR MCD; MGI:96878; LY57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
KW B-cell
FT CONFLICT 133 133 S -> N (IN REF. 2).
FT CONFLICT 147 147 A -> T (IN REF. 2).
FT CONFLICT 148 150 RLA -> GIG (IN REF. 2).
FT CONFLICT 197 198 PP -> AT (IN REF. 2).
FT CONFLICT 391 391 N -> K (IN REF. 2).
FT CONFLICT 444 446 TMD -> SKH (IN REF. 2).
SQ SEQUENCE 457 AA; 50670 MW; 66C93DAFDF9D260 CRC64;


```

Db 428 AsnSerHisGlnHisAsnProLeuValLeuIleAspSerGlnAsn 442
RESULT 12
088504
ID 088504 PRELIMINARY; PRT; 457 AA.
AC 088504;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE B cell linker protein BLNK.
GN LY57.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation.";
RL J. Biol. Chem. 272:27362-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turck C.W., Kurosaki T., Chan A.C.;
RT "BLNK: A Central Linker Protein in B Cell Activation.";
RL Immunity 9:93-103(1998).
DR EMBL: AF068182; AAC40206.1; -.
DR HSSP: P23727; IBFT.
DR MGD: MGI:96878; LY57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS0001; SH2; 1..
SQ SEQUENCE 457 AA; 50803 MW; 66D235796A6C45F0 CRC64;

Alignment Scores:
Pred. No.: 7,79e-17 Length: 457
Score: 287.50 Matches: 106
Percent Similarity: 38.31% Conservative: 53
Best Local Similarity: 25.54% Mismatches: 139
Query Match: 9.40% Indels: 117
Db: 11 Gaps: 16

US-09-856-061-1 (1-1721) x 088504 (1-457)
QY 450 AACAAAGCACTGAGAGATGCTGAGTCCAGCTGCAAGCATGGCCATCAATGAAATTT 509
Db 69 AspSerAspTyrGlnAspProAspGlnHis-----SerAspSerGlnMetTyrVal 85
QY 510 TTACAGCCAGACATATGACGAATCGAATACGAGATACAGCATATTCAGAGATATG 569
Db 86 MetProAlaGluGlnThrGlyAspAspSerTyrGlu----- 97
QY 570 ATGAGAGCTCCCTCTGTACTCTCCCAAGGCTTCTGTCTCCAGTGAAGAGCAAAACGAG 629
Db 98 -----ProProProAla-----GluGlnGlnThrArg 106
QY 630 GATGTGAGG-----ATGACACAGCTGGAAGAGTGGACAAAGCTTACCTTCAAG 677
Db 107 ValValHisProAlaLeuProPheThrArgGlyGlyTyrValAlaAspAsnArgSer----- 124
QY 678 GATGTGAGGCAACCAAGCTTTAA-----GATTCAAATATACAAATAATAACAAAGACT--- 731
Db 125 -----SerGlnArgHisSerProPhePheSerTyrHisThrLeuProSerTyrProSer 141
QY 732 ---CCTTGGCACTCCTCGGCTGCTATATCATCTCTCC-----AAGAG 773
Db 142 TrpProSerAlaValAlaArgValLeuAlaSerThrLeuProAlaProAsnSerLeuGlnLys 161

```

```

QY 774 TACCAACCTTACCCCCAGCACCACCA-----GAGAGAGCAGTGCATCTTCGCT 824
Db 162 ProGln---ValProProLysProLysAspLeuLeuGlnAspGlnAspTyrValVal 180
QY 825 CCAAAGCCCACTTTCACAAAGTCCAGAGGGGGCCAGCAGAGAGT----- 872
Db 181 ProValGlnAspAsnAspGlnAsnTyrTleHisProValGlnSerSerProProAla 200
QY 873 -----GCAAAAGACTTCAGTAGGGTCTTGGAGCAGAGAAGAAATCTCACCACGAGACA 926
Db 201 GlutylAlaProMetValAsnArgSerThrLysProAsnSerSerTyrHisHisMetSer 220
QY 927 AAGCCGAGATCT----- 938
Db 221 ProProGlyThrValAlaGlyArgAsnSerGlyValTrpAspSerLysSerLeuPro 240
QY 939 TCTTGCCCATCATCAACCAACACAGAGAGTCCA----- 977
Db 241 AlaAlaIleProSerProLeuProArgAlaGlyLysLysProAlaThrProLeuLysThrThr 260
QY 978 -----CCTGGCATTGCCAGCTCTTCTTACATAGCCAGAAACACAGATATACACCCAGA 1031
Db 261 ProValProProLeuProAsnAlaSerAsnValLysGlnGlnLysProValProAlaGlu 280
QY 1032 GACCATACAGATGATGATGACGACGACTGCTCTGCTCAGAGATGCCAAGTGCAGCCGAC 1091
Db 281 ArgHisArgLysSer-----SerHis 287
QY 1092 -----AGCCCTGAAATGCTGCCCTTGAACCAACAAACTCGAGAA 1133
Db 288 ArgGlnAspThrValGlnSerProValAlaPheProThrGlnLysProValHisGlnLys 307
QY 1134 CCTGACCCCA----- 1145
Db 308 ProValProLeuProArgPheProGlnAlaGlySerProAlaAlaAspGlyProPheHis 327
QY 1146 -----AAGCTGATGGAAGAGATGCTGTGCGACAAATGATG 1181
Db 328 SerPheProPheAsnLeuThrPheAlaAspGlnGlnGlnLeuGlyLysProTrp 347
QY 1182 TACATTGAGAAATACAGTCCGAGCAGATGGAAGATGTTGTAATGAAGACAAAGAT 1241
Db 348 AlaPhePheAsnLysArgValTyrAsnIleProValArgPheIleGlnAlaThrLysGln 407
QY 1242 TTTCCCTGGGCGACAGACATGAGAAATGAGATGTTGATCTGTGGAAGACATCAT 1421
Db 408 TyrAlaIleGlnGlyLysLysAsnGlyGlnGlyLysPheGlySerValValGlnIleVal 427
QY 1422 GAACACTACACATATTTTCCATCTTCGTAATATGAGGAAGC 1466
Db 428 AsnSerHisGlnHisAsnProLeuValLeuIleAspSerGlnAsn 442

RESULT 13
Q9D413
ID Q9D413 PRELIMINARY; PRT; 297 AA.
AC Q9D413;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 4933424C13RIK protein.
GN 4933424C13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```


QY 564 GATATGATGAGGCTCCCTTCTGTACCTCCCAAGCTTCTGTCTCACTGAGAGACAA 623
Db 184 GlnTyrLeuGlnCysProThrLysIlePro-----AlaLeuHisAspArgMetGln 200
QY 624 ACCAGGATGTGAGTACACAGCTGAGAGAAAGTGGACAAAGCTTCAAGATGTC 683
Db 201 SerLeuAspGlnProThrPheGlnu---GlnGlnGlyAsp----- 212
QY 684 AGAAGCCAAGCTTAAAGATCAATAACACAAAATAACAGACTCTTGCACCT 743
Db 213 -----GlySerAspHisProTyrTyrAsnSerIleProSerLysMet 226
QY 744 CCTGGCCTGCTATCACTCTCCCAAGAGATACCAACCTTACCCCAAGCCACAGAG 803
Db 227 ProProPheGlyGlyPheLeuAspThrArgLeuLysProArgProHisAlaPro----- 244
QY 804 GAGAGCAGTGCATCTTCCGCTCCAAG----- 830
Db 245 ---AspThrAlaGlnPheAlaGlnGlyGlnGlnThrTyrTyrGlnGlnArgHisLeuGly 263
QY 831 CCCACCTTCCAGAA---GTCCAGAGGGGGCCAGCAGAGAGTGCMAAAGACTCAGT 887
Db 264 AspThrPheGlyGlnAspThrGlnGlnThrProLeuArgGlnGlySerSerAsp----- 281
QY 888 AGGTCCTTGGAGCAGAGAGAAATCTCACACCACAGCAAGCCAGAAATCTTGGCCCA 947
Db 282 -----IleTyrSerThrProGlnGlnGlyLysLeuHisValAlaProThrGlnGlnAlaPro 299
QY 948 TCATCAACCAACCAACAGACAGAGAGTCCACCTGCCATTCGCTTCTTACATGCCA 1007
Db 300 ---ThrTyrValAsnThrGlnGlnIleProPro----- 309
QY 1008 GGAAGCAGAGTATACAGCCAGACCAATACAGTACGACGACTGCTGCTCAG 1067
Db 310 -----GlnAlaTyrProAla--- 314
QY 1068 AGATGCCAAGCTGACGACGACCAAGCCTCGA-----ATGCTGCCCTAT 1112
Db 315 ---AlaValSerSerAlaGlnSerSerProThrLysAspLeuPheAspMetLysProPhe 333
QY 1113 GAAACACAAACTCGGAGAAACT----- 1136
Db 334 GlnAspAlaLeuLysAsnGlnProLeuGlnProValLeuSerLysAlaAlaSerValGln 353
QY 1137 -----GACCCACAAAGCCTGATGAGAG-----GATGCTGGCAG 1172
Db 354 CysIleSerProValSerProArgAlaProAspAlaLysMetLeuGlnGlnAla 373
QY 1173 AATGAATGTACATTGGAGATACAGTCCGACGACGACGACGATGATGTTAATGAAGAG 1232
Db 374 GlnThrTyrPheGlnGlnGlnMetSerArgLysGlnAlaGlnGlyLeuLeu----- 390
QY 1233 AACAGAGATGTAATTTTGTCCGAGACTGCTTACAATTCAGAGCAGACCA-- 1289
Db 391 GlnLysAspGlnAspPheLeuValArg-----LysSerThrThrAsnProGly 406
QY 1290 ---TATGTTTGGTGGTGTTTTATGGAGACAGGCTTACAATGTGAATAATCCGTTTCTC 1346
Db 407 SerPheValLeuThrGlnMetHisAsnGlnGlnAlaLysHisLeuLeuValAspPro 426
QY 1347 GAGAGCAATCAACAGTTTGCCCTGGGACAGACTACGAGAAATGAGATGTTGATCT 1406
Db 427 Gln-----GlyThrGlnArgThrLysAspArgValPheAspSer 439
QY 1407 GTGAGACATCATTTGACACTAC-----ACATATTTTCCCATCTGCTAATAGATGG 1460
Db 440 IleSerHisLeuLeuAsnHisHisLeuGlnSerSerLeuProIleVal----- 455
QY 1461 AAAGACAGGCTGACAGCAGAAACAGTGTACTCACCCAGCCTG 1508
Db 456 -----SerAlaGlnSerGlnLeuLys---LeuGlnGlnProVal 467

Search completed: April 21, 2003, 12:28:38
Job time : 130.318 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:22:47 ; Search time 35.4007 Seconds

(without alignments)
1637.367 Million cell updates/sec

Title: US-09-856-061-2

Perfect score: 2316

Sequence: 1 MTSQGNKRRTEKGFGLRQ.....QPLPLARLLLTQYSSQALHE 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2316	100.0	435	22	AAAB81070
2	1316.5	56.8	428	23	AAE22609
3	1316.5	56.8	443	23	AAE22608
4	1132	48.9	376	22	AAAB81071
5	1109.5	47.9	353	23	AAE22610
6	326	14.1	456	22	AAAB93406
7	319	13.8	449	19	AAW59866
8	319	13.8	456	19	AAW59865
9	287.5	12.4	457	19	AAW59867
10	260	11.2	95	20	AAW70587

11	213	9.2	84	21	AAAG03152
12	212	9.2	46	22	ABAB8147
13	212	9.2	46	22	ABAB23354
14	212	9.2	46	22	AAW58783
15	212	9.2	46	22	AAW71296
16	212	9.2	46	22	AAW18973
17	212	9.2	46	22	AAW31575
18	212	9.2	46	22	ABG41096
19	185.5	8.0	503	22	ABG22854
20	165.5	7.1	474	18	AAW15253
21	165.5	7.1	594	18	AAW15256
22	150.5	6.5	743	22	AAW79738
23	144.5	6.2	474	18	AAW39089
24	144.5	6.2	474	18	AAW39087
25	142.5	6.2	474	18	AAW39084
26	142.5	6.2	474	18	AAW39092
27	141.5	6.1	474	18	AAW39085
28	141.5	6.1	474	18	AAW39088
29	141.5	6.1	474	18	AAW39090
30	141.5	6.1	474	18	AAW39091
31	141.5	6.1	474	18	AAW39095
32	141.5	6.1	474	18	AAW39096
33	141.5	6.1	474	18	AAW33246
34	141.5	6.1	474	18	AAW15257
35	141	6.1	728	22	AAW78754
36	140.5	6.1	474	18	AAW39094
37	140	6.0	469	17	AAW97242
38	140	6.0	540	22	AAW40461
39	140	6.0	640	22	AAW38675
40	138.5	6.0	474	18	AAW39086
41	138.5	6.0	474	18	AAW39093
42	138.5	6.0	923	22	AAE11773
43	137.5	5.9	594	18	AAW65631
44	137.5	5.9	594	18	AAW15260
45	134.5	5.8	787	22	AAW30364

ALIGNMENTS

RESULT 1
AAB81070
ID AAB81070 standard; Protein: 435 AA.
XX
AC AAB81070;
XX
DT 25-JUN-2001 (first entry)
XX
DE Murine mast cell-specific signal transduction protein.
XX
KW Mast cell; signal transduction; mouse; allergic disease.
XX
OS Mus musculus.
XX
PN JP3146204-B1.
XX
PD 12-MAR-2001.
XX
PE 17-SEP-1999; 99JP-0263778.
XX
PR 17-SEP-1999; 99JP-0263778.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR WPI: 2001-310022/33.
XX
DR N-PSDB: AAF86139.
XX
PT Mast cell-specific signal-transduction molecule, useful for screening
XX
PT therapeutic compounds for treating allergies, is specifically expressed
XX
PS by mouse mast cell -
XX
Claim 1; Page 7-8; 12pp; Japanese.

CC This sequence represents a murine mast cell-specific signal transduction
 CC protein. The invention includes the cDNA and protein sequences of the
 CC mast cell-specific signal transduction molecule and an expression vector
 CC containing the polynucleotide sequence. The coding sequence of the signal
 CC transduction protein can be used for screening therapeutic compounds
 CC which will be useful for treating allergic diseases.

XX Sequence 435 AA:

Query Match 100.0%; Score 2316; DB 22; Length 435;
 Best Local Similarity 100.0%; Pred. No. 4.3e-194;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTSQGNKRTTKEGFDLRFQNVSLKNSRSPSLSSAKGRCAVLEPLPDHRRNLAVPGG 60
 DB 1 MTSQGNKRTTKEGFDLRFQNVSLKNSRSPSLSSAKGRCAVLEPLPDHRRNLAVPGG 60
 QY 61 EKCNNSNDYEDPEFQLLKAMPKMLIPAPRTOESEYADRYQDMMEAPLLPPKASVST 120
 DB 61 EKCNNSNDYEDPEFQLLKAMPKMLIPAPRTOESEYADRYQDMMEAPLLPPKASVST 120
 QY 121 ERQTRDVMTQLEVDKPTFKDVRSQRFKGYTKINKTLPPLPPPAITLPKKYQPLPPA 180
 DB 121 ERQTRDVMTQLEVDKPTFKDVRSQRFKGYTKINKTLPPLPPPAITLPKKYQPLPPA 180
 QY 181 PPESSAIFAPKPTPEVQGRPRORSADKDFSVILGAEEESHQTKPESSCPSSNONTOKS 240
 DB 181 PPESSAIFAPKPTPEVQGRPRORSADKDFSVILGAEEESHQTKPESSCPSSNONTOKS 240
 QY 241 PPAIASSSYMPCGKHSIOARDHTGSMOHCPOACOAASHSPMLYEYENTNSEKPPPTKPD 300
 DB 241 PPAIASSSYMPCGKHSIOARDHTGSMOHCPOACOAASHSPMLYEYENTNSEKPPPTKPD 300
 QY 301 EKDVQWQNEWYIGEYSRQAVEDVLAKENKDGTFVLVDCSTKSAEPYVLVFEYGNKYVNVK 360
 DB 301 EKDVQWQNEWYIGEYSRQAVEDVLAKENKDGTFVLVDCSTKSAEPYVLVFEYGNKYVNVK 360
 QY 361 IFFLESNQOAFALGTGRGNEMFDSVEDIIEHTYTPILLIDGKDKAARRKOCYLTQPLPL 420
 DB 361 IFFLESNQOAFALGTGRGNEMFDSVEDIIEHTYTPILLIDGKDKAARRKOCYLTQPLPL 420
 QY 421 ARLLTOYSSOALHE 435
 DB 421 ARLLTOYSSOALHE 435

RESULT 2

ID AAE22609 standard; Protein: 428 AA.

XX AAE22609;

DT 26-JUL-2002 (first entry)

XX Human MIST splice variant protein from clone #7.

KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;

KW lymphocytes; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;

OS Homo sapiens.

PN WO200226986-A2.

PD 04-APR-2002.

PE 28-SEP-2001; 2001WO-US30593.

PR 29-SEP-2000; 2000US-237030P.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Perez-Villar JT, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;

XX WPI: 2002-372126/40.
 DR N-PSDB: AAD35801.

PT New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma

PS Claim 11; Fig 5; 171pp; English.

CC The present invention relates to novel mast cell immunoreceptor signal
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.
 CC MIST sequences of the invention are useful for preventing, treating or
 CC ameliorating a medical condition in mammalian subject. They are useful
 CC for treating an immune disorder involving hyperactivity of B- or T-
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
 CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
 CC as targets for therapeutic intervention in immune cell disorders and
 CC inflammatory indications, for diagnosis and/or screening of disorders
 CC or diseases associated with expression of MIST, for screening for
 CC antagonists or inhibitors of the interaction of MIST with cellular
 CC signalling components. They are used in assays that detect activation
 CC or induction of various B and T-cell-related neoplasms or cancers.
 CC Sequences of the invention are also used in gene therapy. The present
 CC sequence is human MIST splice variant protein from clone #7.

XX Sequence 428 AA:

Query Match 56.8%; Score 1316.5; DB 23; Length 428;
 Best Local Similarity 62.1%; Pred. No. 1.2e-106;
 Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

QY 1 MTSQGNKRTTKEGFDLRFQNVSLKNSRSPSLSSAKGRCAVLEPLPDHRRNLAVPGG 60
 DB 1 MTSQGNKRTTKEGFDLRFQNVSLKNSRSPSLSSAKGRCAVLEPLPDHRRNLAVPGG 60
 QY 61 EKCNNSNDYEDPEFQLLKAMPKMLIPAPRTOESEYADRYQDMMEAPLLPPKASVST 120
 DB 61 AKGSHDDDDYDPELMEETWQSIKILPARPIKESYADTHFKVAMOTPLDTRTSISI 120
 QY 121 ERQTRDVMTQLEVDKPTFKDVRSQRFKGYTKINKTLPPLPPPAITLPKKYQPLPPA 180
 DB 121 GQPTWNTQ-TRLEKVDKPTISKDVRSQNKIKGASVKNKPIPLPPPLITLPKKYQPLPP 178
 QY 181 PPESSAIFAPKPTPEVQGRPRORSADKDFSVILGAEEESHQTKPESSCPSSNONTOKS 240
 DB 179 EPSSRRPPLSQRHPTPEVQGRPSQISLDLSEVLEAEKVPKRNKPESTHLENONTOEI 238
 QY 241 PPAIASSSYMPCGKHSIOARDHTGSMOHCPOACOAASHSP--PMLYEYENTNSEKPPPTK 298
 DB 239 PLAISSSFTTNSHVSQNRDRGOMQPCSPQRCPPASCSPHENILPYKYSWMPRPKR 298
 QY 299 PDEKDVQWQNEWYIGEYSRQAVEDVLAKENKDGTFVLVDCSTKSAEPYVLVFEYGNKYVNV 358
 DB 299 SDRDVQWQNEWYIGEYSRQAVEDVLAKENKDGTFVLVDCSTKSAEPYVLVFEYGNKYVNV 358
 QY 359 VKIFFLESNQOAFALGTGRGNEMFDSVEDIIEHTYTPILLIDGKDKAARRKOCYLTQPL 417
 DB 359 VKIFFLESNQOAFALGTGRGNEMFDSVEDIIEHTYTPILLIDGKDKTGVHRRKOCYLTQPL 418
 QY 418 LPLARLL 425
 DB 419 LPLTRHLL 426

RESULT 3

ID AAE22608 standard; Protein: 443 AA.

XX AAE22608;

DT 26-JUL-2002 (first entry)

```

DE Human MIST protein #1.
XX
XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
XX inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 84 /label= Tyrosine_phosphorylation_binding_site
FT Modified-site 111 /label= Tyrosine_phosphorylation_binding_site
FT Domain 306..311 /label= SH3_binding_proline-rich_motif
FT Domain 324..407 /label= SH2_domain
FT Domain /label= SH2_domain
XX
PN MO200226986-A2.
PD 04-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US30593.
XX
PR 29-SEP-2000; 2000US-237030P.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Perez-Villar J, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX
DR WPI: 2002-372126/40.
XX
DR N-PSDB; AAD55800.
XX
PT New isolated mast cell immunoreceptor signal transducer polypeptide,
XX useful for treating immune disorder involving hyperactivity of B- or
XX T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX
PS Claim 11; Fig 3; 171pp; English.
XX
XX The present invention relates to novel mast cell immunoreceptor signal
XX transducer (MIST) proteins and polynucleotides encoding such proteins.
XX MIST sequences of the invention are useful for preventing, treating or
XX ameliorating a medical condition in mammalian subject. They are useful
XX for treating an immune disorder involving hyperactivity of B- or T-
XX lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
XX as targets for therapeutic intervention in immune cell disorders and
XX inflammatory indications, for diagnosis and/or screening of disorders
XX or diseases associated with expression of MIST, for screening for
XX antagonists or inhibitors of the interaction of MIST with cellular
XX signalling components. They are used in assays that detect activation
XX or induction of various B and T-cell-related neoplasms or cancers.
XX Sequences of the invention are also used in gene therapy. The present
XX sequence is human MIST protein.
XX
SQ Sequence 443 AA:
XX
Query Match 56.8%; Score 1316.5; DB 23; Length 443;
Best Local Similarity 62.1%; Pred. No. 1.3e-106;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;
XX
QY 1 MTSQGNKRTTEGFGDRLFQVNSLLKNRSPSLSSAKGRCRAVLEPLDPHRRNLGAGVGG 60
DB 16 MNRQGNKRTTEGFGDRLFQVNSLLKNRSPSLSSAKGRCRAVLEPLDPHRRNLGAGVGG 75
XX
QY 61 EKCSNNDYEDPEFOLKAMPMSKILPAPLOESEYADTRYFODMEAPLLPPKASVST 120
DB 76 AKGSHDDYDDPELMEETWOSIKLPAPRIKESYADTRYFODMEAPLLDPFTSISI 135
XX
QY 121 ERQTRDAMTOLDEYDKTFQDVNSQKRGFKYKINKTPLPPPPAATLTKKIQPLPA 180
DB 136 GQPTWNTQ-TFLERVDKPKISKDVRSQNIKGDASVAKKNIKPLPPPPAATLTKKIQPLPP- 193

```

```

QY 181 PRESSATYAPKPTPEVQGRORSAKDESRVLAGAEEESHQFPKSSCPSSNONTOKS 240
DB 194 EPSSSRPPLSQRTTPEVQMPQISLRDLSEVLEAEKPHNQRKPESTHLENQTOEL 253
XX
QY 241 PPAIASSSYWPKGHSIOARDHTGSMOCPAORCOAAASHP--RMLPYENTNSEKDDPTK 298
DB 254 PLAISSSFTTSMHSVQNRHRGGMQPCSPQRQPPASCPHENILPYKTSMRPFPRK 313
XX
QY 299 PDEKDWQNEWTYGEYSROADEVLYLAKENKDGTFVLVRODSTKSAEPYLVVYGGKYVN 358
DB 314 SDRKDVQHNEMWTYGEYSRQAVEAFMKENKDGSLVRDCSTKSEPPYLVAVEENKYVN 373
XX
QY 359 VKIRPLESNOQFALGTGLGENEMEDVEDIEHYTFEPILLIDGDK-AARRKOCYLTOP 417
DB 374 VKIRFLERNQFALGTGLGDEKDEVEDIEHYTKNFPILLIDGDKYGRKQCHLTOP 433
XX
QY 418 LPLARLL 425
DB 434 LPLRLHL 441
XX
RESULT 4
AAB81071
ID AAB81071 standard; Protein; 376 AA.
XX
AC AAB81071;
XX
DT 25-JUN-2001 (first entry)
XX
DE Human mast cell-specific immunoreceptor signal transducer.
XX
KW Mast cell; signal transduction; human; allergic disease; MIST;
XX mast cell-specific immunoreceptor signal transducer.
XX
OS Homo sapiens.
XX
PN JP3146204-BL.
XX
PD 12-MAR-2001.
XX
PF 17-SEP-1999; 99JP-0263778.
XX
PR 17-SEP-1999; 99JP-0263778.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR WPI: 2001-310022/33.
XX
DR N-PSDB; AAF6140.
XX
PT Mast cell-specific signal-transduction molecule, useful for screening
XX therapeutic compounds for treating allergies, is specifically expressed
XX by mouse mast cell -
XX
PS Example 1; Page 10-11; 12pp; Japanese.
XX
XX This sequence represents human MIST (mast cell-specific immunoreceptor
XX signal transducer). The invention relates to cDNA encoding a murine mast
XX cell-specific signal transduction protein. Included in the invention are
XX cDNA and protein sequences of the mast cell-specific signal transduction
XX molecule and an expression vector containing the polynucleotide sequence.
XX The coding sequence of the signal transduction protein can be used for
XX screening therapeutic compounds which will be useful for treating
XX allergic diseases.
XX
SQ Sequence 376 AA:
XX
Query Match 48.9%; Score 1132; DB 22; Length 376;
Best Local Similarity 60.9%; Pred. No. 1.5e-90;
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;
XX
QY 19 FQVNSLLKNRSPSLSSAKGRCRAVLEPLDPHRRNLGAGVGGKCSNNDYEDPEFOLK 78
DB 1 FQNSLIPKRSMPRINSATGQYORMNKPLDMERNFAVLDGAKGSHDDYDDPELMEE 60

```

```

OY 79 AMFSMKILPAPRIQSEYADTRYFQDMMEAPLLPPKASVSTERQTDVMTQLEEDVP 138
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 TMOQSIILPAPRIKESYADTRYFQDMMEAPLLPPKASVSTERQTDVMTQLEEDVP 119
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 139 TERDVASQRFKGGKTKYIKKTLPPRPRAITLPPKYYQPLPPAPPESSAYFAPKPTPEY 198
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 ISHDVRSQNIKGDASVYRKRIPLPPRPRLITLPPKYYQPLPP-EPSSRPPLSQRHHTPEY 178
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 199 OSGPRRSKAKDFSRVYLGAEESHQTKPESSCPSSNQNTQKSPRALASSYMPGKHSIQ 258
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 QGMPQSILKRDSEVLEAKVPHNQRPSTHLENNQTEIPLAIISSSFTTSMHSYQ 238
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 259 RDHTGSMOHCPORCQAAASHSP--RMLPYENTNSEKPPDPTKPEKDVQNMWYIGEYSR 316
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 RDHRGSMQPCSPQRCOPRASCSPHENILPYKYSWRPPRPKRSKRDQVQNMWYIGEYSR 298
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 317 QAVEDVLMKENDGTFVLVDCSTKSKAEPLYLVFYGNKYVVKIRFLESNOQFALGTG 376
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 QAVEEAFMKENDGSLVDCSTKSKKEPYVLAVFYENKYNVKIRFLERNOQFALGTG 358
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 377 RGNEMFDSVEDITEHY 392
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 RGEKFDSEVEDITEHY 374
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
AAE22610
ID AAE22610 standard; Protein: 353 AA.
AC AAE22610;
XX
XX 26-JUL-2002 (first entry)
XX
XX Human MIST splice variant protein from clone #12.
DE
XX
XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX
XX Homo sapiens.
OS
XX
XX WO200226986-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US30593.
XX
XX 29-SEP-2000; 2000US-237030P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
PA
XX
XX Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
PI
XX
XX WPI: 2002-372126/40.
XX
XX N-PSDB: AAD35802.
XX
XX New isolated mast cell immunoreceptor signal transducer polypeptide,
XX useful for treating immune disorder involving hyperactivity of B- or
XX T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma -
XX
XX Claim 11; Fig 8; 171pp; English.
XX
XX The present invention relates to novel mast cell immunoreceptor signal
XX transducer (MIST) proteins and polynucleotides encoding such proteins.
XX MIST sequences of the invention are useful for preventing, treating or
XX ameliorating a medical condition in mammalian subject. They are useful
XX for treating an immune disorder involving hyperactivity of B- or T-
XX lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
XX as targets for therapeutic intervention in immune cell disorders and
XX inflammatory indications, for diagnosis and/or screening of disorders
XX or diseases associated with expression of MIST, for screening for

```

```

CC antagonists or inhibitors of the interaction of MIST with cellular
CC signalling components. They are used in assays that detect activation
CC or induction of various B and T-cell-related neoplasms or cancers.
CC Sequences of the invention are also used in gene therapy. The present
CC sequence is human MIST splice variant protein from clone #12.
XX
XX
SQ Sequence 353 AA;
Query Match 47.9%; Score 1109.5; DB 23; Length 353;
Best Local Similarity 64.4%; Pred. No. 1.2e-88;
Matches 226; Conservative 31; Mismatches 89; Indels 5; Gaps 4;
OY 78 KAMPQSKILPAPRIQSEYADTRYFQDMMEAPLLPPKASVSTERQTDVMTQLEEDVP 137
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3.ETMOQSIILPAPRIKESYADTRYFQDMMEAPLLPPKASVSTERQTDVMTQLEEDVP 61
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 138 PTFKDVRSQRFKGGKTKYIKKTLPPRPRAITLPPKYYQPLPPAPPESSAYFAPKPTPE 197
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 PISKDVRSQNIKGDASVYRKRIPLPPRPRLITLPPKYYQPLPP-EPSSRPPLSQRHHTPE 120
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 198 VQGRPRQSAKQFYSVYLGAEESHQTKPESSCPSSNQNTQKSPRALASSYMPGKHSIQ 257
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 VQRMPSQISLRLDSEVLEAKVPHNQRPSTHLENNQTEIPLAIISSSFTTSMHSYQ 180
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 258 ARDHTGSMOHCPORCQAAASHSP--RMLPYENTNSEKPPDPTKPEKDVQNMWYIGEYSR 315
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 NRDRGSMQPCSPQRCOPRASCSPHENILPYKYSWRPPRPKRSKRDQVQNMWYIGEYSR 240
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 316 ROAVEDVLMKENDGTFVLVDCSTKSKAEPLYLVFYGNKYVVKIRFLESNOQFALGTG 375
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 ROAVEEAFMKENDGSLVDCSTKSKKEPYVLAVFYENKYNVKIRFLERNOQFALGTG 300
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 376 RGNEMFDSVEDITEHYTFPLILIDGKD-AARKQCYLQPLPLALL 425
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 LRGEKFDSEVEDITEHYKNPFLILIDGKDGVRKQCHLTQPLPLTRHLL 351
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
AAB93406
ID AAB93406 standard; Protein: 456 AA.
AC AAB93406;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:12602.
DE
XX
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.
KW
XX
XX Homo sapiens.
OS
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the

```


PT full-length cDNAs -
 XX Claim 8; SEQ ID 12602; 2537bp + CD ROM; English.
 PS
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 XX Sequence 456 AA;
 SQ
 Query Match 14.1%; Score 326; DB 22; Length 456;
 Best Local Similarity 24.6%; Pred. No. 6.6e-20;
 Matches 99; Conservative 50; Mismatches 141; Indels 112; Gaps 9;
 YY 64 NSNNVDYEDPEFOLKAMPKSKILPAPRIQSEVADTYFDQMEAP---LLLPKASVST 120
 DB 91 NADSYEPPEVE---QETRPVHPALPFARGEYIDNRSSO--RHSPFSKTLSPKSPWPS 144
 YY 121 EROTRDVRMTQLEEVDRKTFKDVRSQRFKGYTKINKTLPPLPAPITLPPKYOPLPPA 180
 DB 145 EKARLTSTLPALATALOKRPV-----PPKPRGLLEADYVVPV 182
 YY 181 PPESSAYFAKPTTFPEVQKPRORSKADFSVLGAEEESHQTPSSCPSSQNT--- 237
 DB 183 EDNDENYIHPTESSPPPEKAPM-----VNRSTKPMNSTPASPSPGTAAG 226
 YY 238 -----QKSPPAIASS-----SYMPGHSIQARHTQSMQ 266
 DB 227 RNSGAWETKSPPPAAPSPLPRACKKPTTPLKTPPVASQANASVCGEKPIPAERHRGSS 286
 YY 267 HCPAORCOAASHSPMLPYENTNSEKRPD-----TKPDEK 302
 DB 287 -----KQEAQVQSPVFPFPAOKOIHOKRIPILPRTTEGNTLVGDLPSFSNSTISQDA 339
 YY 303 DVMQNEWYIGESRQAVEDVLMKENKDGTLVRCSTKSKAEPYLVVFGKNVYANKIR 362
 DB 340 GVLCKPMPYAGACDRKSAEALHRSNKDGSFLIRKSSGHDSCOPYTLVVFENKRYNIPVR 399
 YY 363 FLESNOQFALGTGLRGNEMPDSVEDIIEHTYTPILLIDGD 404
 DB 400 FIEATKOYALGRKKNGEYFGSVAEITRNHOSHPLVLIDSON 441
 RESULT 7
 AAWS9866
 ID AAWS9866 standard; Protein; 449 AA.
 XX
 AC AAWS9866;
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE Amino acid sequence of the human BLNK-2 protein.
 XX
 KW Human; BLNK-2; B cells linker protein-2; apoptosis; Grb2; PLC-gamma;

KW SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; Sos;
 KW ras pathway; GDP; GTP; calcium pathway; antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO9832852-A1.
 XX
 PD 30-JUL-1998.
 XX
 PF 23-JAN-1998; 98WO-US01394.
 XX
 PR 17-MAR-1997; 97US-0819013.
 XX
 PR 24-JAN-1997; 97US-0788322.
 XX
 PA (UNITW) UNIV WASHINGTON.
 XX
 PI Chan AC, Fu C;
 XX
 DR WPI: 1998-427948/36.
 XX
 DR N-PSDB; AAV41902.
 XX
 PT Human B cell linker proteins - useful in the treatment of diseases
 XX involving increased or decreased apoptosis
 XX
 PS Disclosure; Fig 3; 56pp; English.
 XX
 CC This is the amino acid sequence of the human BLNK-2 (B cells linker
 CC protein-2) protein, used the treatment of diseases involving the
 CC increase and decrease of apoptosis, in the method of the invention.
 CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
 CC intracellular calcium levels and Grb2 is an adapter molecule
 CC containing two SH3 domains that mediate its interaction with the
 CC guanine nucleotide exchange factor. Son of Sevenless (Sos) which in
 CC turn activates the ras pathway by facilitating the exchange of GDP for
 CC GTP on the ras molecule. Activation of both ras and calcium pathways
 CC are required for efficient B cell antigen receptor function. BLNK
 CC binding proteins can be used to identify BLNK proteins in a target
 CC sample.
 XX
 XX Sequence 449 AA;
 SQ
 Query Match 13.8%; Score 319; DB 19; Length 449;
 Best Local Similarity 25.0%; Pred. No. 2.7e-19;
 Matches 99; Conservative 58; Mismatches 139; Indels 100; Gaps 12;
 YY 64 NSNNVDYEDPEFOLKAMPKSKILPAPRIQSEVADTYFDQMEAP---LLLPKASVST 120
 DB 84 NADSYEPPEVE---QETRPVHPALPFARGEYIDNRSSO--RHSPFSKTLSPKSPWPS 137
 YY 121 EROTRDVRMTQLEEVDRKTFKDVRSQRFKGYTKINKTLPPLPAPITLPPKYOPLPPA 180
 DB 138 EKARLTSTLPALATALOKRPV-----PPKPRGLLEADYVVPV 175
 YY 181 PPESSAYFAKPTTFPEVQKPRORSKADFSVLGAEEESHQTPSSCPSSQNT--- 237
 DB 176 EDNDENYIHPTESSPPPEKAPM-----VNRSTKPMNSTPASPSPGTAAG 219
 YY 238 -----QKSPPAIASSYMP--GKH-----SIQARHTQSM--QKCPAORCOAA- 276
 DB 220 RNSGAWETKSPPPAAPSPLPRACKKPTTPLKTPPVASQANASVCGEKPIPAERHRGSS 278
 YY 277 -----ASHSPMLPYENTNSEKRPD-----TKPDEKVMQNE 308
 DB 279 HROEAQVQSPVFPFPAOKOIHOKRIPILPRTTEGNTLVGDLPSFSNSTISQDAVGLCKP 338
 YY 309 WYIGESRQAVEDVLMKENKDGTLVRCSTKSKAEPYLVVFGKNVYANKIRFLESNO 368
 DB 339 WYAGACDRKSAEALHRSNKDGSFLIRKSSGHDSCOPYTLVVFENKRYNIPVRIEATK 398
 YY 369 QFALGTGLRGNEMPDSVEDIIEHTYTPILLIDGD 404
 DB 399 QYALGRKKNGEYFGSVAEITRNHOSHPLVLIDSON 434

Query Match	Best Local Similarity	13.8%	Score 319,	DB 19,	Length 456;
Matches 99;	Conservative 58;	Mismatches 139;	Indels 100;	Gaps 12;	
64	NSNNYEDPPEFLKAWPSMKILPARPIQESYADTRFQDMMEAP---	LLPPRASVST	120		
91	NADSDYEPPEPVE---OETRPVHPLPARGETIDNRSSQ--RHSPPSKTLPSKPSWS	144			
121	EQRTDVAWMTGLEVDKPTFKDVRNSQRFKGYTKINKTPTLPPPPATILPKKYQPLPFA	180			
145	EKARLTSTYLPALTAOKRQVY-----PPKPKGLDEADYVVPV	182			
181	PPESASAFAPAKPTPEYQGRQGRQSAKDPSRVLGAEEESHQOTPESSCPSSNONT---	237			
183	EDNENENTYHPRESSPPPEKAPM-----VNSTKPSNSTPASPPTASG	226			

```

QY 238 -----OKSPRALASSYVM--CKH-----SIGARHITGSM---QHCMPARCOAA- 276
Db 227 RNSGAMETKSPSP--AAPSPLCPACKPTPTPLKTPPVASQOMASSVCECKPIPERHNGSS 285
QY 277 ----ASHSPRLPYENTSEKPPD-----TKPEKDWQNE 308
Db 286 HQGEAVOSPYRPPAKQIHOXKPIPLRPTREGNPTVDOGPIRPSNSTISEQEGVLCRP 345
QY 309 WIGEYSKQAVEDVLMKENKDGTEFLVNRDCSTKSKAEFYVAVPEYGNKVYVWKITFLESNO 368
Db 346 WYAGACDCKRSKAEMLHRSNKNKGSEFLIRKSSGHDSKQPYTLVFEFNKRKVYNIIPREIATK 405
QY 369 QFALGTGLRGNEMEDSVDDILEHTTYPRILLIDSKD 404
Db 406 QYALGRKNGEYFQSVALLIRNHQSHPLVLDSON 441

```

Query Match	12.4%	Score 287.5	DB 19	Length 457
Best Local Similarity	25.5%	Pred. No. 1.6e-16		
Matches 106; Conservative	53;	Mismatches 139;	Indels 117;	Gaps 16;

CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNA. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

XX Sequence 84 AA;

Query Match Best Local Similarity 9.2%; Score 213; DB 21; Length 84;

Matches 41; Conservative 17; Mismatches 26; Indels 0; Gaps 0;

QY 1 MTSQGNKRTKRGFDLRFQNVSLKNSRPSLSAKGRCRAVLEPLDHRRLAGVPG 60

DB 1 MNRQGNKRTKRGSDNLKFNFSLPKNSRPSRINSATGQYRMKPLLDWERNFAVLDG 60

QY 61 EKCSNNDYEDPEFQLKAPSMK 84

DB 61 AKGSHDXXDDPELRMEETQSIK 84

RESULT 12

ABB38147 ID ABB38147 standard; Peptide: 46 AA.

AC ABB38147;

DT 04-FEB-2002 (first entry)

DE Peptide #5653 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN W0200157277-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001MO-US00669.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-483447/52.

DR Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver.

XX Claim 27; SEQ ID NO 30782; 639pp + sequence listing; English.

PS The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 46 AA;

XX Query Match 9.2%; Score 212; DB 22; Length 46;

Best Local Similarity 89.1%; Pred. No. 2.7e-11; Matches 41; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 329 DGFVLVRCSTKSKAEPYLVVFGNKVYVVKIRFLESNOQFALGT 374

DB 1 DGSFLVRCSTKSKAEPYLVVFGNKVYVVKIRFLESNOQFALGT 46

RESULT 13

ABB23354 ID ABB23354 standard; Protein: 46 AA.

XX ABB23354;

DT 23-JAN-2002 (first entry)

DE Protein #5353 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

OS Homo sapiens.

PN W0200157274-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001MO-US00666.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488899/53.

DR Single exon nucleic acid probes for analyzing gene expression in human

PT hearts.

XX Claim 15; SEQ ID NO 25124; 530pp; English.

PS The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease.

CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 46 AA;

XX Query Match 9.2%; Score 212; DB 22; Length 46;

XX Best Local Similarity 89.1%; Pred. No. 2.7e-11;

XX Matches 41; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 329 DGFVLVRCSTKSKAEPYLVVFGNKVYVVKIRFLESNOQFALGT 374

DB 1 DGSFLVRCSTKSKAEPYLVVFGNKVYVVKIRFLESNOQFALGT 46

Db 91 NADSEYEPPEV-----QETRPVHPALPFARGEYIDNRSSQ--RHSPPEFKTLPSKSPWS 144
QY 121 EROTRDVMTOLEEVYDKPTFKVRSQRKGFYTKINTPLPPRPATLTPKKYQPLPRA 180
Db 145 EKARLTSLPALALOKQY-----PPKPKGLLEADYVAV 182
QY 181 PPESSAFAKPRPEVQGRFORSAKDFSVLGAEEESHQTPRESSCSNONOT--- 237
Db 183 EDNDENYIHPTESSSPPEKAPM-----VNRSTKPNSTASPPGTASG 226
QY 238 -----OKSPPAIASSIMP--GKH-----SIQARDHTGSM---QHCPAQRCQA- 276
Db 227 RNSGAMETKSPBP--AAPPLPRAGKPPYPLKTPPVASQONASSVCEKPIPAERHRGSS 285
QY 277 ---ASHSPRLPYENTSEKDP-----TKPDEKDYQNE 308
Db 286 HROEVQSPVPPPAKQKHPKIPLPRTTEGPNPTVDGPLPFSSNSGSISSQEGVLCKP 345
QY 309 WYGEYSQAVEDVLMKENDGTFLVRDCTSKAEPYLVVYFYNKYKIRPLESNO 368
Db 346 WYAGACDRKSAEALHRSNKDGSFLIRKSGHDSKQPTLVVFFMKRYNIPVRIEATK 405
QY 369 QPALGTGRGEMDVSVDIIEHTYFFILLIDGKD 404
Db 406 QYALGRKKNGEYFGSAEIIIRNHQHSPLVLIDSON 441

RESULT 2

US-08-729-416C-1
Sequence 1, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-1

Query Match 7.1%; Score 165.5; DB 3; Length 474;
Best Local Similarity 23.2%; Pred. No. 5.5e-07;
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;
QY 13 CFDDLRFQNVSLKNSWPSLSSAKGRCRAVLPLRPHRNRLAGVPGEGKCNNDY--- 69

Db 94 GKSNIQFAGMSLSTIATSLNLTTPDSKQI--ANHHMRISIFASGDP--DTTDYVAX 149
QY 70 --EDP---EFOLLKAMPMSKILTPARPIOESEYADTRYFODMEAPLLPPKASVSTERO 123
Db 150 VAKDENVNRACHILLECCGG---LAQDVIGSIGQAFELFKQYLQCPFTIP---ALHDMQ 203
QY 124 TRDVMTOLEEVYDKPTFKVRSQRKGFYTKINTPLPPRPATLTPKKYQPLPRAPE 183
Db 204 SLDEPWT-EEGD-----GSDHPYNSIPSKMPPPGFLDTRLRKPRHAP-- 247
QY 184 ESSAFAK-----PTEPE-VORGPFORSAKDFSVLGAEEESHQTPRESSCP 231
Db 248 -DIAQFAKEQTYOGRHLGDTFGEDMQOTPLRQSSD---ISTPEGKLHVAFTGAP 302
QY 232 SSNONTOKSPPAIASSSYMPGKHSIQARDHTGSMOHCAPQRCQAASHSPR-----MLPY 286
Db 303 -TYVNTQIIP-----QAWPA--AVSSAESPPKKDLFDMPKF 336
QY 287 ENTSEK-----DPTKPEK---DYMONEWTIGEYSROAVEDVLMKE 326
Db 337 EDALKNQPLGVLKSAASVECI SPVSPRAPPAKMLELQAEIETWYQEGMSRKEAEGL- 393
QY 327 NKDGTFLVRDCTSKAEP--YVLVYFYNKYKIRPLESNOQFALGTGLRGEMFDS 384
Db 394 EKDDGFLVR-----KSTTPGGSFVLGHNQAKHLLVDPE-----GITRTQDRVFD 442
QY 385 VEDIIEHY--TYFPILLIDGKDKAARRKQCYLTQPL 418
Db 443 ISHLINHLESPLIV-----SAGSELC-LQDPV 470

RESULT 3

US-08-729-416C-7
Sequence 7, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-7

Query Match 7.1%; Score 165.5; DB 3; Length 594;
 Best Local Similarity 23.2%; Pred No. 7.6e-07;
 Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;

13 GGGDLRFQVNSLLKNSWPSLSAKGRCRAVLEPLDHRNLAGVGGKCSNNNDY--- 69
 214 GKSNIQFAGMSISLTSTASLNLRTPDSKQII--ANHHKRSISFASGSDP--DTTDYVAY 269
 70 --EDP-----EFLQKAMPMSKILIPARFQSEADRTYQDMMEAFLILPKRASVSTENQ 123
 270 VAKDPNRRACHILECCDG--LAQDVISIGQAFELRKYLCPTKIP--ALHDMQ 323
 124 TRDVRMTOLEVDKPTFKDVRSQRFKFKYTKINKTLPPLPPRPAITLPKKYOLPPAPPE 183
 324 SLDEPTE-EEGD-----GSDHPYNSITPKMPRPGGLDRLRKRPAP-- 367
 184 ESSAYAPK-----PTFPE-VQGRPROSAKDFSRVLAGAEESHQTKPSSCP 231
 368 -DTAQFAGKEQTYGGRHLGDTFGEEDMOOTPLRGSSD---IYSTPEGLHVAPTGEAP 422
 232 SSNQNTQKSPPAIASSYMPGKHSIQARDHTGSMOHCAPQCOAASHSPR-----MLPY 286
 423 -TYVNTQQLP-----QAMPA-AVSSAESSPRKLFDMPKP 456
 287 ENTNSEKP-----DPTKDEK--DVMQNEWYIGESQAVEDYLMKE 326
 457 EDLKNQPLGPRVLSKAASVEICISPVSPRAPDAKMLELAETWYQGMSEKAEGL--- 513
 327 NKGGTLVRDCSTKSKAEP--YVLVVFYGNKYNNKIRPLESNOQFALGTGLGNEMFDS 364
 514 EKGGDELVR---KSTNTPGSFVLTGMHNGOAKHLLLVPE-----GTRTRKDRVFD 562
 385 VEDIIEY--TYPIILLIDGDKAARRKOCYLOPL 418
 563 ISHLINHLESLPIV-----SAGSELC-LOQPV 590

RESULT 4

US-08-807-342B-2
 ; Sequence 2, Application US/08807342B
 ; Patent No. 6077686

GENERAL INFORMATION:
 APPLICANT: Der, Channing
 APPLICANT: O'Bryan, John P.
 APPLICANT: Pawson, Anthony
 TITLE OF INVENTION: No. 6077686el SHC Proteins
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Howson and Howson
 STREET: Spring House Corporate Cntr., P.O. Box 457
 CITY: Spring House
 STATE: Pennsylvania
 COUNTRY: USA

ZIP: 19477
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/807,342B
 FILING DATE: 28-FEB-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/012,516
 FILING DATE: 29-FEB-1996

ATTORNEY/AGENT INFORMATION:
 NAME: Bak, Mary E.
 REGISTRATION NUMBER: 31,215
 REFERENCE/DOCKET NUMBER: MTS2USA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-540-9200
 TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 474 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-807-342B-2

Query Match 6.1%; Score 141.5; DB 3; Length 474;
 Best Local Similarity 23.0%; Pred. No. 8.7e-05;
 Matches 101; Conservative 66; Mismatches 177; Indels 95; Gaps 24;

13 GGGDLRFQVNSLLKNSWPSLSAKGRCRAVLEPLDHRNLAGVGGKCSNNNDY--- 69
 94 GKSNIQFAGMSISLTSTASLNLRTPDSKQII--ANHHKRSISFASGSDP--DTTDYVAY 149
 70 --EDP-----EFLQKAMPMSKILIPARFQSEADRTYQDMMEAFLILPKRASVSTENQ 123
 150 VAKDPNRRACHILECCDG--LAQDVISIGQAFELRKYLCPSKVP--ALQDMQ 203
 124 TRDVRMTOLEVDKPTFKDVRSQRFKFKYTKINKTLPPLPPRPAITLPKKYOLPPAPPE 183
 204 SLDEPTE-EEGD-----DHPYNSVPTKMPRPGGLDRLRKRPAP-- 247
 184 ESSAYAPK-----PTFPE-VQGRPROSAKDFSRVLAGAEESHQTKPSSCP 231
 248 -EAAQFAGKEQTYGGRHLGDTFGEEDMOAPTRGSLDIYST--ABGKTH--WVPGERP 302
 232 SSNQNTQKSP-----AIASSYMPGKHSIQARDHTGSMOHCAPQCOAASHSPMLPY 286
 303 -TYVNTQPPRPQWPAATSTESSPPRKLFDMPFEDALRNQPLSLSKAASVEICISPV 361
 287 ENTNSEKPPTKDEK--DVMQNEWYIGESQAVEDYLMKENKGTFLVRDCSTSKA 343
 362 -----TPRADAKMLELAETWYQGMSEKAEGL---REGDELVR---KSTT 406
 344 EP--YVLVVFYGNKYNNKIRPLESNOQFALGTGLGNEMFDSVEDIIEY--TYPIILL 399
 407 NPGSFVLTGMHNGOAKHLLLVPE-----GTRTRKDRVFDISHLINHLESLPIV- 458
 400 IDGDKAARRKOCYLOPL 418
 459 -----SAGSELC-LOQPV 470

RESULT 5

US-08-353-550-6
 ; Sequence 6, Application US/08353550
 ; Patent No. 5744313

GENERAL INFORMATION:
 APPLICANT: Williams, Lewis T.
 APPLICANT: Cavanaugh, William M.
 TITLE OF INVENTION: No. 5744313el Protein Domain Which Binds
 Tyrosine Phosphorylated Proteins
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Kourie and Crew
 STREET: One Market Plaza, Stewart Tower, Suite 2000
 CITY: San Francisco
 STATE: California
 COUNTRY: USA

ZIP: 94105
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/353,550
 FILING DATE: 09-DEC-1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Dow, Karen B.

```

APPLICATION NUMBER: US/08/551,687
FILING DATE: 01-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/353,550
FILING DATE: 09-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Matthew B.
REGISTRATION NUMBER: P39,787
REFERENCE/DOCKET NUMBER: 2307K-5731
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-551-687-6

Query Match          6.08; Score 140; DB 2; Length 469;
Best Local Similarity 22.38; Pred. No. 0.00012;
Matches 96; Conservative 55; Mismatches 185; Indels 94; Gaps 18;

13 GFGDLRFQNVSLILNKRSPSLSSAKRCRAVLEPLDHRRLAGVGGKCSNNDY--- 69
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
98 GKSLRFAGMSISLHISIDGLSVPRATROYI--ANHHMPISFASGSD--TDMIDYAY 153
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
70 --EDPEQ---LKKAPSMKILPARPIQES--EYADTRYQDMMEAP-LLLPKASVST 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
154 VAKPINORACHILECEGEL---AQSLISTVGQAFELRFKQYLIHSPKVALPPERLAGP 209
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
121 ERQTRDVRMTGLE-----EVDKPTFKDVSQRKFGKRYKINTPPRPPRATLT 170
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
210 EESMGGEEDLEHNYNSIGKEPRPLGLVDSLRLALQPCALRALDGGPSPSLDACS 269
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
171 PKKQPLPRAPRESSAFAKPTTFPEVQ---RGRPRSAKDFSRVLGAEEESHQTKPE 227
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
270 FMDVSGTGTAPRGD-----YQADARGP---PDHEEHLVYVNTQGLDAPRPE 313
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
228 SSCSSNQNGKSPRALASSYMPGKHSIQARHDHTGSMQHCPADRCQAAASHSRMLPYE 287
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
314 DS-----PKKDLFDMRPREDALK--LHECSVAAGVTAPRLPLE 349
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
288 NTNSEKPD---PTKPKDEKDVQNMENYIGEVSRQAVEDVLMKENNGDTFLVDCSTKSAE 344
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
350 DQMPSPETRRAPVAFTEQQLQRPWYHGRMSRAAENL---RADGDFLVADSVT--NPG 404
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
345 PYVLVYGVKKYVAVKTRFLESNQFALGCLGKRNEMFDSVEDIIEHTYFPILLIDGKD 404
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
405 QYVLGTHMAQPKILLLVDE-----GVVTRKDVLFESISHLIDHH-----LQNGQP 451
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
405 KAARRKOCYL 414
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
452 IVAAESLHL 461
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 7
US-08-729-416C-11
Sequence 11, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA

```

```

; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-729-416C-11

Query Match          5.9%; Score 137.5; DB 3; Length 474;
Best Local Similarity 21.7%; Pred. No. 0.0002;
Matches 95; Conservative 70; Mismatches 180; Indels 93; Gaps 23;

QY 13 GFGDLRFQNSLTKNRSPLSSAKGRCRAVLEPLPDHRRNLAVPGGKCNNSNDY--- 69
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 94 GKSLQFAGMSISLTISTASINLRTPDSKQITIS--NHHMSISFASGDDP--DTTDVYAY 149
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 70 --EDP-----EFOLKAMPMSKILPARPIQSEYADTRYFDMMEARPLLPKASVSTERQ 123
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 150 VAKPVPNRACHILLECCDG--LAQDVIGSIGQAFELRFKQYLOCPSKIP--ALQDRMQ 203
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 124 TRDVRMTQLEVDKPTFKDVSQRFGFKYTKIKTLPPLPPRAITLPPKKYQPLPPAPPE 183
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 204 SLDEPWTE-EEGDGP-----DHPYNSVNPKNMPPGCGFLDARKARPHAP-- 247
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 184 ESSAYFAKPTF-----PEVQGRPRQSAKDFSRVLAEEESHQTKRESSCPS 232
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 248 DAAQFSGEQTYIGRHLGDAFGEDWQAPTRQSLD---IYSTPEGKAHWV-PVGETP- 302
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 233 SNQNTQKSP-----AIASSSYMPGKHSIOARDHTGSMQHCPOAQAASHSPRLPYE 287
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 303 TYVMTQPPVPQVMPAATSTESSPRKDLFDMPKPFEDALRNQPLGVLSKASVEICISPV- 361
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 288 NTSEKPDPTKPEDEK---DVQWQNEWTIGEYSRQAVEDVLMKENDGTFVLVDCSTKSAE 344
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 362 -----TPRAPDAKMLELNAEPYQGEKSRKEAALL--QEDGDELVR---KSTTN 407
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 345 P--YVLVVFYGNKYNNKIRFLESNQGFALGTGLNGEMPDVEDIIEHY--TYFPILLI 400
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 408 PGSEFVLTMHNGQAKHLLVDP-----CTVTKDKRVFDSISHLITYHLESSLPIY-- 458
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 401 DGKDKAARRKQCYLTQPL 418
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 459 -----SAGSELC-LRQPV 470
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 8
US-08-729-416C-17
; Sequence 17, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.M.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 594 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-729-416C-17

Query Match          5.9%; Score 137.5; DB 3; Length 594;
Best Local Similarity 21.7%; Pred. No. 0.00028;
Matches 95; Conservative 70; Mismatches 180; Indels 93; Gaps 23;

QY 13 GFGDLRFQNSLTKNRSPLSSAKGRCRAVLEPLPDHRRNLAVPGGKCNNSNDY--- 69
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 214 GKSLQFAGMSISLTISTASINLRTPDSKQITIS--NHHMSISFASGDDP--DTTDVYAY 269
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 70 --EDP-----EFOLKAMPMSKILPARPIQSEYADTRYFDMMEARPLLPKASVSTERQ 123
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 270 VAKPVPNRACHILLECCDG--LAQDVIGSIGQAFELRFKQYLOCPSKIP--ALQDRMQ 323
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 124 TRDVRMTQLEVDKPTFKDVSQRFGFKYTKIKTLPPLPPRAITLPPKKYQPLPPAPPE 183
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 324 SLDEPWTE-EEGDGP-----DHPYNSVNPKNMPPGCGFLDARKARPHAP-- 367
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 184 ESSAYFAKPTF-----PEVQGRPRQSAKDFSRVLAEEESHQTKRESSCPS 232
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 368 DAAQFSGEQTYIGRHLGDAFGEDWQAPTRQSLD---IYSTPEGKAHWV-PVGETP- 422
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 233 SNQNTQKSP-----AIASSSYMPGKHSIOARDHTGSMQHCPOAQAASHSPRLPYE 287
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 423 TYVMTQPPVPQVMPAATSTESSPRKDLFDMPKPFEDALRNQPLGVLSKASVEICISPV- 481
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 288 NTSEKPDPTKPEDEK---DVQWQNEWTIGEYSRQAVEDVLMKENDGTFVLVDCSTKSAE 344
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 482 -----TPRAPDAKMLELNAEPYQGEKSRKEAALL--QEDGDELVR---KSTTN 527
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 345 P--YVLVVFYGNKYNNKIRFLESNQGFALGTGLNGEMPDVEDIIEHY--TYFPILLI 400
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 528 PGSEFVLTMHNGQAKHLLVDP-----CTVTKDKRVFDSISHLITYHLESSLPIY-- 578
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 401 DGKDKAARRKQCYLTQPL 418
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 579 -----SAGSELC-LRQPV 590
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 9
US-08-426-509A-4
; Sequence 4, Application US/08426509A
; Patent No. 6326469
; GENERAL INFORMATION:

```

APPLICANT: Ullrich, Axel
 APPLICANT: Gishlitzky, Mikhail
 APPLICANT: Sures, Irman G.
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
 TITLE OF INVENTION: TYROSINE KINASES
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York,
 STATE: NY
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: PASTSEQ Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/426,509A
 FILING DATE: 21-APR-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/232,545
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7683-0074-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 675 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: No. 6326469e
 US-08-426-509A-4

Query Match 5.7%; Score 133; DB 4; Length 675;
 Best Local Similarity 24.0%; Pred. No. 0.00088;
 Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKRYOPLPAPPESSAYFAKPTPE-VQRGPR-----QSAKDFRYLGA--- 216
 DB 141 PGCLTMEAYANLHTAVNEKHRV-----PTFPDRVLKIRAVPVLMKDAFSSSTLLAQYDN 196
 QY 217 EESHQTKPESSCPSSNQNTOKSPAIASS-----SYMPG-----KHSIOAR 259
 DB 197 ESKKNYGSQPPSSSTSLAQYDSNSKKIYGSQPNFMQYIIPREDFPDMMQVRLKSSSSSE 256
 QY 260 DHGSMOHCRAQOQAASHSPRLPYENTNSEKPDPTKPEKDVQWQNEWYIGEYSROAV 319
 DB 257 DVASSNOK-----ERNVYHTTSKISWEFPSSSE-----EENLDYDWMFAGNISRSQS 306
 QY 320 EDVLMKRNKDGTLVPRDCSTKSKAEPYLVLF---YGNKYVNVKIRPLESNOQFALGTGL 376
 DB 307 EQLLRQKKGKGAFAVNRN---SSQYGMVTVSLFSKAVNDKKGTVKHHVHTNAENKLL--YL 361
 QY 377 RGNEMFDSVEDIIEHYTY 394
 DB 362 AENYCFDSIPKLTHYHGH 379

RESULT 10
 PCT-US95-05008-4
 Sequence 4, Application PC/TUS9505008
 GENERAL INFORMATION:
 APPLICANT: Sugen, Inc.
 APPLICANT: 515 Calveston Drive

APPLICANT: Redwood City, California 94063-4720
 APPLICANT: United States of America
 APPLICANT: Wissenschaften E.V.
 APPLICANT: Hofgarten Str. 2
 APPLICANT: Munchen 80539
 APPLICANT: Germany
 TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
 KINASES
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/05008
 FILING DATE: 24-APR-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/232,545
 FILING DATE: 22-APR-1994
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7683-074
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)790-9090
 TELEFAX: (212)869-9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 675 amino acids.
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 PCT-US95-05008-4

Query Match 5.7%; Score 133; DB 5; Length 675;
 Best Local Similarity 24.0%; Pred. No. 0.00088;
 Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKRYOPLPAPPESSAYFAKPTPE-VQRGPR-----QSAKDFRYLGA--- 216
 DB 141 PGCLTMEAYANLHTAVNEKHRV-----PTFPDRVLKIRAVPVLMKDAFSSSTLLAQYDN 196
 QY 217 EESHQTKPESSCPSSNQNTOKSPAIASS-----SYMPG-----KHSIOAR 259
 DB 197 ESKKNYGSQPPSSSTSLAQYDSNSKKIYGSQPNFMQYIIPREDFPDMMQVRLKSSSSSE 256
 QY 260 DHGSMOHCRAQOQAASHSPRLPYENTNSEKPDPTKPEKDVQWQNEWYIGEYSROAV 319
 DB 257 DVASSNOK-----ERNVYHTTSKISWEFPSSSE-----EENLDYDWMFAGNISRSQS 306
 QY 320 EDVLMKRNKDGTLVPRDCSTKSKAEPYLVLF---YGNKYVNVKIRPLESNOQFALGTGL 376
 DB 307 EQLLRQKKGKGAFAVNRN---SSQYGMVTVSLFSKAVNDKKGTVKHHVHTNAENKLL--YL 361
 QY 377 RGNEMFDSVEDIIEHYTY 394
 DB 362 AENYCFDSIPKLTHYHGH 379

RESULT 11
 US-08-807-342B-4

```

Query Match          5.7%; Score 132.5; DB 3; length 550;
Best Local Similarity 21.1%; Pred. No. 0.00073;
Matches 103; Conservative 65; Mismatches 188; Indels 133; Gaps 21

QY      13 GFGDIRFQNVSLILKNRSMPSLSSAGRCRAVLEPLRDHRRLNAGVPGKCKNSNDY----- 69
      106 GKSNIREFGMSISVSNIVDGLNLSYPATRQIL--ANHHMQSSIFASGSD--TDMTDYAV 161
QY      70 ---EDDEPQ-----LLKAWPSMKLLPRAPVQSESYADTRYQDMMEAPLLIPKASVSTEQ 123
      162 VAKDDINORACHILECCCEGLAOSVISTVGOA-----FELRFKOYLHSPRAVPE-- 212
QY      124 TRDVMTQLEEV-----DKPTFKDV--RSORFKGKYYKINKTP----- 160
      213 -----RLTLEELIEMAGDDDAADNHYNSIPGKEPLGLGVDSRLAVTQCALATLGGGLG 268
QY      161 ---LPPRAPIATLPKKYYQQLPRAPRESSAYFAPKTFPPVYQ---RPPRORSAKDSRVLG 215
      269 GMTFPWRAARGPLPMDMGSGAAPPDGG-----YVQADARPH----- 305
QY      216 AEEESHNOTKPESSCPSSNQNTQKSPRAIASSYMPGKHSIQARDHTGSMOCHPACRQA 275
      306 -DYEEHLVYNNQGLDAVELEDTAEAPLQFEDS---PKKLEFDMRPPEDALK---LHACSV 358
QY      276 AASHSPRLPLVIENTSEKPD---PTKPRDEKDVQNEWMTIGEYSKQAVEDVILKENKDGTF 332
      359 AAGITTAASPLEDDQWPSPTTRAPAPLAPTEEOQLROEPMWYHGRRSSRAAEKTL---RADGDF 415
QY      333 LVPRDCKSKAEPLYLVAFYGNKYNVAKIRPLESNQAFGLGGLNGENMEDSVEDYIIIEHY 392
      416 LVYRDSVT---NPGQYVLTGMHAGQPHHLLVDP-----GVYRKTKDVLFESISHLTDIH 466

```

Db 232 -ALALTDGSPSLRDLACSLPMDVSGTAPPGDG-----YVQADARGP--- 296
 Oy 206 SAKDFSRVIGAEESHSNHTKPESSCPSSNONOTKSPPAIASSYMPGKHSIOARDHTGSM 265
 Db 297 --PDHEHLVYVNTQGLDAPEPEDS-----PKKDLDMREEDAL 333
 Oy 266 QHCPRQCOAAASHSPRLPYENTNSEKPD---PTKPEKDVQWQEWYIGEYSRQAVEDV 322
 Db 334 K---LHECSVAGVTAAPLPLEDCWPSPTTRAPYAPTEEQLROEPWYHGRMSRRAERM 390
 Oy 323 LMKENKDGFLVRDCST 339
 Db 391 L---RADGDFLVRDSVT 404

RESULT 13 US-07-906-349A-5

; Sequence 5, Application US/07906349A
 ; Patent No. 5434064
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlessinger, Joseph
 ; APPLICANT: Skolnik, Edward Y.
 ; APPLICANT: Margolis, Benjamin L.
 ; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
 ; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AN
 ; NUMBER OF INVENTION: TARGET PROTEINS
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Browdy and Neimark
 ; STREET: 419 Seventh Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/906,349A
 ; FILING DATE: 30-JUN-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/643,237
 ; FILING DATE: 18-JAN-1991
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-628-5197
 ; TELEFAX: 202-737-3528
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 724 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-906-349A-5

Query Match 5.5%; Score 128; DB 1; Length 724;
 Best Local Similarity 20.6%; Pred. No. 0.0028;
 Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;

Oy 81 PSMKILPARPIQSESYADTRVFQDMMEAPLLLPKAS----- 117
 Db 85 PTPKRPRPRPLVPAP-GSSKTEADVEQOALTLPLDAEQFAPPDIAAPLLIKLVEAIEKKG 143
 Oy 118 -----VSTERQTRVVRMTGLEVDKPTFK-----DVRSQRFKGFYTKINKPTPLPPR 165
 Db 144 LECSTLYRTQSSSLAEIRQLDCTPSVDLEMIDVHLADAFKRYL-----LDLPN 195
 Oy 166 PAITLPKKYQPLPPAPRESSAYFAP-----KTPPEVQGRQPRQASAKDFSRVIGABE 218
 Db 196 PVIIPAIVSEMISSLAPEVQSSSEETIQLLKLIRSPSIHQIWLTLQYLKHKFKL----- 250

Oy 219 ESHQTKPESSCPSSNONOTKSPPAIASSYMPGKHSIOARDHTGSM----- 265
 Db 251 -----SQTSKULLNARVLSEIFSPMLPFRFSAASSDNENLIKVEIILLISTEWN 299
 Oy 266 --QHCPRQCOAAASHSPRLPYENTNSEKPDPTFK-----DEKDVQWQEWYIGEX 314
 Db 300 EROPAPA-----LP-----PKPRPTTVANNGNMNNNSIQNAEYWGDI 338
 Oy 315 SROAVEDVLMKENKDGFLVRDCSTPKSAPRYVLVFPYGNKYVVKIRFLESNOOFALGT 374
 Db 339 SREYVNEKL-RDTADGTILVADASTKMGD--YTLFLRKGCH--KKLIKIFHRDKYGRSD 394
 Oy 375 GLRGNEFMFDSVEDIIEHY 392
 Db 395 PL---TFSSVELINHY 408

RESULT 14 US-08-167-035-2

; Sequence 2, Application US/08167035
 ; Patent No. 5618691
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlessinger, Joseph
 ; APPLICANT: Skolnik, Edward Y.
 ; APPLICANT: Margolis, Benjamin L.
 ; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
 ; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: 10036-2711
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/167,035
 ; FILING DATE: 16-DEC-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cotruzz1, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7683-062
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 724 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-167-035-2

Query Match 5.5%; Score 128; DB 1; Length 724;
 Best Local Similarity 20.6%; Pred. No. 0.0028;
 Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;

Oy 81 PSMKILPARPIQSESYADTRVFQDMMEAPLLLPKAS----- 117
 Db 85 PTPKRPRPRPLVPAP-GSSKTEADVEQOALTLPLDAEQFAPPDIAAPLLIKLVEAIEKKG 143
 Oy 118 -----VSTERQTRVVRMTGLEVDKPTFK-----DVRSQRFKGFYTKINKPTPLPPR 165
 Db 144 LECSTLYRTQSSSLAEIRQLDCTPSVDLEMIDVHLADAFKRYL-----LDLPN 195

QY 166 PAITLPKKYQPLPPAPPESSAYFAP-----KPTPEVQRRPRORSARDFSRVGAEE 218
 DB 196 PVIPAAVYSMISLAPEVQSSEYIQLKKLIRSPSPHGYWMLTQYLKHFRL----- 250
 QY 219 ESHHQTKPESSCPSSNONTKSPPAIASSSYMKGHSIQARDHTGSM----- 265
 DB 251 -----SQTSSKNLNLNARVLSIFSPMLFRFSAASSDNTENLIKVIETLISTEWN 299
 QY 266 --QHCPAQRCAAAASHSPRLPYENTNSEKPDPTK-----DEKDVQWQEWYIGEX 314
 DB 300 ERQPPAP-----LP-----PKPKPTTVANGNKNMMSIQNAEWYIGDI 338
 QY 315 SROAEDVLMKENDGTFVLRDCSTKSKAPRYLVYFGKKNVYVKIRPLESNOQFALGT 374
 DB 339 SREVNNEKL-RDTADGTFVLRDASTKMGD-YTLTLRKGNN-NKLKIKIFHRDQKYGFS 394
 QY 375 GLRGNEKPDSEVEDIIEHY 392
 DB 395 PL-----TFSSVVELINHY 408

RESULT 15

US-08-208-887A-2
 ; Sequence 2, Application US/08208887A
 ; Patent No. 5677421
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlessinger, Joseph
 ; APPLICANT: Skolnick, Edward Y.
 ; APPLICANT: Margolis, Benjamin L.
 ; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
 ; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
 ; NUMBER OF SEQUENCES: 51
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; City: New York
 ; STATE: New York
 ; COUNTRY: 10036-2711
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/208,887A
 ; FILING DATE: 11-MAR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7683-063
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 724 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-208-887A-2

Query Match 5.5%; Score 128; DB 1; Length 724;

Best Local Similarity 20.6%; Pred. No. 0.0028;

Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;

QY 81 PSMKILPAPPIQSESYADTRYFQDMMEAPLLIPKAS----- 117
 DB 85 PTPKRPRLPVPAD-GSSKTEADVEQOALTLPDLAEQFAPPDIAPLLIKLVEAIKKKG 143

QY 118 -----VSTRQTRDVRAMTOLEEVDKPTK-----DVRQRRKGFYTKINKTLPBPR 165
 DB 144 LECSTLYRTQSSSNLAELRLQDLDDPPSDLEMDIVHVLADAKRRL-----LDLPN 195
 QY 166 PAITLPKKYQPLPPAPPESSAYFAP-----KPTPEVQRRPRORSARDFSRVGAEE 218
 DB 196 PVIPAAVYSMISLAPEVQSSEYIQLKKLIRSPSPHGYWMLTQYLKHFRL----- 250
 QY 219 ESHHQTKPESSCPSSNONTKSPPAIASSSYMKGHSIQARDHTGSM----- 265
 DB 251 -----SQTSSKNLNLNARVLSIFSPMLFRFSAASSDNTENLIKVIETLISTEWN 299
 QY 266 --QHCPAQRCAAAASHSPRLPYENTNSEKPDPTK-----DEKDVQWQEWYIGEX 314
 DB 300 ERQPPAP-----LP-----PKPKPTTVANGNKNMMSIQNAEWYIGDI 338
 QY 315 SROAEDVLMKENDGTFVLRDCSTKSKAPRYLVYFGKKNVYVKIRPLESNOQFALGT 374
 DB 339 SREVNNEKL-RDTADGTFVLRDASTKMGD-YTLTLRKGNN-NKLKIKIFHRDQKYGFS 394
 QY 375 GLRGNEKPDSEVEDIIEHY 392
 DB 395 PL-----TFSSVVELINHY 408

Search completed: April 21, 2003, 12:39:04
 Job time: 15.4094 secs

Qy	121	ERORDVMTQLEEDKRTFEDVDSQOFKGGKYYIKITPLPPRAVITLTPKYOLPPA	160
		121	121
Dp	121	GOPTMTNQ-TLLEKYDKFISNDVSSQMIKGASVAKKKITPLPPRPLITLTPKTYOLPP-	178
Qy	181	PPRESSAYFADKPPTPEVQNGPRQORSKADFSRYLGAEEESHQTKPESSCPSSNOMQKS	240
Dp	179	EPESRRPEPLSQRHPTPEVQNRPSQISLRDLESEVLEAKVPHNQRPPESTHLEQNQTQEI	238
Qy	241	PPATASSSYMGGKISIOARHGTGSMQCPOROCQAASHSP--RMLEPYMTNSEKPRPTK	298
Dp	239	PLATSSSSFTTSSNHSYQNRHGRGMOQCSFORCOPRASCSPHEHILPYKTTSMRPPPKR	298
Qy	299	PDEKQWQNEYITGEYSQOAEVDLMEKNDGTFVLVDCSTKSAAEYVLVEYGNKVYN	358
Dp	299	SDRDVQYHNEWYIEEYSQOAEVEFMENKNDGSEFLVDCSTKSKEEYVLVAEYENKVYN	358
Qy	359	VKIRFLESNQGFALGTGHRGNEMDSEVEDIIEHTYTPILLIDGKKD--AARRQCVLTOP	417
Dp	359	VKIRFLENGQFALGTGRDPEKDSVEDIIEHTKNEPILLIDGKKDTGVHRRQCHLTOP	418
Qy	418	LPLARLLI 425	
Dp	419	LPLTRHLL 426	

```

RESULT 2
US-09-966-955A-2
; Sequence 2, Application US/09966955A
; Patent No. US20020155563a1
; GENERAL INFORMATION:

```

```

APPLICANT: Perez-Villar, Juan J.
APPLICANT: Chang, Han
APPLICANT: Yang, Wen-Pin
APPLICANT: Wu, Yuli
APPLICANT: Whitney, Gena S.
APPLICANT: Kanner, Steven B.
TITLE OF INVENTION: Identification and Cloning of a Full-length Human
TITLE OF INVENTION: Clink-related Gene, MIST-(Mast Cell Immunorecept)
FILE REFERENCE: 3053-413US1
CURRENT APPLICATION NUMBER: US/09/966,955A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/237030
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 443
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
OTHER INFORMATION: HUMAN FULL-LENGTH MIST CDNA CLONE #8, TRANSLATED
OTHER INFORMATION: AMINO ACID SEQUENCE
IS-09-966-955A-2

```

Query Match	56.8%	Score	1316.5	DB	9	Length	443
Best Local Similarity	62.1%	Pred. No.	5.7e-85				
Matches	266	Conservative	46	Mismatches	111	Indels	5
				Gaps			4

```

QY 1 MTSOGKRRTRKEFGDRLRFONVSLKRRSPSLSSAKGRCAVLEPLPDHRRNLGAVPG 60
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 16 MNRGSRKRTTKEGSDNLKFONFSLPKRRSPRINSATGOYGRNNKPLLDMERNAVLDG 75
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 EKCSNNDDYEDDPQLLKAMPMSKILPARIQSEVADTRYFODMMAPLLLPKASVST 120
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 AKGSHSDDDYDPELRMEETQSIKILPARIKESEVADTRYFKVADTPPLDTRISIL 135
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 ERQTRVRMTQLEEVDPKTRKDVRSQRFKFKTKINKTLPPLPPRAITLPKKYQPLPPA 180
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 GQETPLWNTQ-TRLEVRVDPKPSIKDVRSQNIKEDASVRRKKITPLPPRPRLITPKKYQPLPP- 193
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 PPESSAVYAPKRTFPVQGPGRQSRSAKDDSRVLGAEEESHQDTPEBSSCPSSNONTQKS 240
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	134	EPBSSRPLSQRHFFPEVQMPQSISLRLDSEVLAEKVPINQRPPESTHLENOQEI	2533
Qy	241	PRAIASSYMGKHSIQARDHTGSMQHCPRACQOAAASHSP--RMLYEYNTNSEKRPDTK	2898
Db	254	PLAISSSFTTSMHSVQNRDHRGGMQPCSPQRCQPRASCSPHEIILRYKXTSMRPFPKR	3133
Qy	259	PDEKQWQNEWTIGETSHQAVEDVLMKENKGTGLVYRDCSTKSKAEPPYLVLTGKNVYN	3568
Db	314	SDRKQVQNEWTIGETSHQAVEAEFMKENKGSFLVYRDCSTKSKAEPPYLVAVFENKVVYN	3733
Qy	359	VKIRELESNOQFALGTGIRGNEMDPVEDIIEHTYFPIILIDDKDK--AARRKQCYLTOP	4173
Db	374	VKIRELENGQFALGTGIRGEKRDSDVEDIIEHTKKNPILIDDKDKYGVHNRQCHLTQ	4333
Qy	418	LPLARLL 425	
Db	434	LPLTRHLL 441	

RESULT 3
US-09-966-955A-6
; Sequence 6, Application US/09966955A
; Patent No. US20020155563A1

```

APPLICANT: Perez-Villar, Juan J
APPLICANT: Chang, Han
APPLICANT: Yang, Wen-Pin
APPLICANT: Wu, Yuli
APPLICANT: Whitney, Gena S.
TITLE OF INVENTION: Identification and Cloning of a Full-length Human
TITLE OF INVENTION: Cink-related Gene, MIST (Mast Cell Immunoreceptor
TITLE OF INVENTION: Signal Transducer)
FILE REFERENCE: 3053-4130S1
CURRENT APPLICATION NUMBER: US/09/366,955A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/237030
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 353
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #12, TRANSLATED
OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-6

```

Query Match	47.9%	Score 1109.5	DB 9	Length 353
Best Local Similarity	64.4%	Pred. No. 1.5e-70		
Matches 226	Conservative 31	Mismatches 89	Indels 5	Gaps 4

Qy	78	KAMSMILPAPRQIESTYADTRFQOMMEAPLLPRKASVSTROTRIDVAMTOLEBYDK	13
Db	3	ETWOSILILPRPKESEYADTHFKAMDPRLDTRISISTQIPWMMNO-TLERDK	61
Qy	138	PTFDVRSQRFKGFKKYKINTPLPPRPATILTPKTKYQPLPAPPESSAFAPKPTPE	197
Db	62	PISDVRSQNIKGDASVRKNNILPLPPRPLTLTKKQYPLP-EPESRPLPSRHITFE	120
Qy	198	VQRPQRQSAKDFSRVLGAEEESHQTKPESSCSSNMONTKSPRAIASSYMFGKHSIO	257
Db	121	VQRPQSISLDELEVLEAEKVPINQOKPESTHLEMONQOELPLAASSSEFTTSSHSVQ	180
Qy	258	ARDHTSGMOHCPAORCOAASHSP--BMLPEANTNSEKPDPTKDEKDVOMONEYIEBS	315
Db	181	NROHGRGMOQCSPORCOPRASCSEHNIILPKYTSMSRPPPKSKDRKDVOQHNEYIEBS	240
Qy	316	ROAVEDVLAKENKCDTFLVRDCSTKSKAEPRVLVVPYGNKYVYKIRFLESNOOFALGTG	379
Db	241	ROAEERFMENKESGLSLVDCSTKSKSEEBYVLAVFENKYVYKIRFLELRNOOFALGTG	300

PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 675
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
US-09-977-260-4

Query Match 5.7%; Score 133; DB 9; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.13;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKKYQPLPPAPPESSAYFAKPTPE-VORGPR-----QSAKDSRYLGA--- 216
DB 141 PGCTLMEAYANLHTAVNEKHRV---PTFPDRVLKIPRAVPVLMKDPASSSTTLAQYDN 196
QY 217 EESHHOTKPRESSCPSSNONTKSPPAIASS-----SYMPG-----KHSIQAR 259
DB 197 ESKKNYSGPPSSSTSLAQYDSNKKIYGSQPNFMQYIPREDPDMQVRLKSSSSSE 256
QY 260 DHTGSMQHCAPQCAAAASHSPRLPYENTNSEKPDPTKDEKDYWQNEWYIGEYSRQAV 319
DB 257 DVASSNOK-----ERNVHTTSKISWEPESSSE---EENLDYDWMFAGNISRSQS 306
QY 320 EDVLKKNKDGTFILVRDCTKSKAEPYLVVF---YGNKYVNVKIRPLESNOQFALGTGL 376
DB 307 EQLLRQKKGEGAFWVRN---SSQVGMVTVLSFSAVNDKKGTVKHYHVTNAENKL--YL 361
QY 377 RGNEMFDSVEDIIEHYTY 394
DB 362 AENYCFDSIPKLHYHGH 379

RESULT 7
US-09-977-261-4
Sequence 4, Application US/09977261
Publication No. US20030054527A1
GENERAL INFORMATION:
APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1259
CURRENT APPLICATION NUMBER: US/09/977,261
PRIOR FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 675
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
OTHER INFORMATION: Kinase 2
US-09-977-261-4

Query Match 5.7%; Score 133; DB 9; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.13;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKKYQPLPPAPPESSAYFAKPTPE-VORGPR-----QSAKDSRYLGA--- 216
DB 141 PGCTLMEAYANLHTAVNEKHRV---PTFPDRVLKIPRAVPVLMKDPASSSTTLAQYDN 196
QY 217 EESHHOTKPRESSCPSSNONTKSPPAIASS-----SYMPG-----KHSIQAR 259
DB 197 ESKKNYSGPPSSSTSLAQYDSNKKIYGSQPNFMQYIPREDPDMQVRLKSSSSSE 256

QY 260 DHTGSMQHCAPQCAAAASHSPRLPYENTNSEKPDPTKDEKDYWQNEWYIGEYSRQAV 319
DB 257 DVASSNOK-----ERNVHTTSKISWEPESSSE---EENLDYDWMFAGNISRSQS 306
QY 320 EDVLKKNKDGTFILVRDCTKSKAEPYLVVF---YGNKYVNVKIRPLESNOQFALGTGL 376
DB 307 EQLLRQKKGEGAFWVRN---SSQVGMVTVLSFSAVNDKKGTVKHYHVTNAENKL--YL 361
QY 377 RGNEMFDSVEDIIEHYTY 394
DB 362 AENYCFDSIPKLHYHGH 379

RESULT 8
US-09-977-269-4
Sequence 4, Application US/09977269
Patent No. US20020082037A1
GENERAL INFORMATION:
APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977,269
PRIOR FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 675
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
OTHER INFORMATION: Kinase 2
US-09-977-269-4

Query Match 5.7%; Score 133; DB 10; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.13;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKKYQPLPPAPPESSAYFAKPTPE-VORGPR-----QSAKDSRYLGA--- 216
DB 141 PGCTLMEAYANLHTAVNEKHRV---PTFPDRVLKIPRAVPVLMKDPASSSTTLAQYDN 196
QY 217 EESHHOTKPRESSCPSSNONTKSPPAIASS-----SYMPG-----KHSIQAR 259
DB 197 ESKKNYSGPPSSSTSLAQYDSNKKIYGSQPNFMQYIPREDPDMQVRLKSSSSSE 256
QY 260 DHTGSMQHCAPQCAAAASHSPRLPYENTNSEKPDPTKDEKDYWQNEWYIGEYSRQAV 319
DB 257 DVASSNOK-----ERNVHTTSKISWEPESSSE---EENLDYDWMFAGNISRSQS 306
QY 320 EDVLKKNKDGTFILVRDCTKSKAEPYLVVF---YGNKYVNVKIRPLESNOQFALGTGL 376
DB 307 EQLLRQKKGEGAFWVRN---SSQVGMVTVLSFSAVNDKKGTVKHYHVTNAENKL--YL 361
QY 377 RGNEMFDSVEDIIEHYTY 394
DB 362 AENYCFDSIPKLHYHGH 379

RESULT 9
US-10-081-980B-1
Sequence 1, Application US/10081980B
Publication No. US20030041337A1
GENERAL INFORMATION:
APPLICANT: Pfizer Inc.
APPLICANT: Gibbs, E. Michael
APPLICANT: McNeish, John D.
TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form

QY	38	GRRAVLEELPDHR--RNLGAVPGGEKCNKNNDD--PEFOLIKAMPREKILTPARIGES	94
Db	78	GRKR-ISPPTKPRPRPLPVAPGSSKTEADTEGOALPLDLAEQAFAPDVAAPLLIKL	136
QY	95	EYADRRYQDMMWEAPLLPPKASVSTERQTR----DYRMTQLEEVDRKPTFKDVSQRFK	149
Db	137	EAIEKK-----GLECSTLKRQSS--SNRRELQOLDCCAASVDLEMDIVHLADA-----FK	187
QY	150	GFKYTKINKTLP-PP-----PPRAITLKKYQLPAPPE--ESSAYFAPK	192
Db	188	--RYLADLPNPVIVPAAVYNEMMSLAQELQSPEDCIQLLKLRLPNIPIHQCLTLQYDLK	245
QY	193	PTFFVQVQPRQSAKD-----TSRYL-----GAEEESHQOTPRESSCPSSNNT	237
Db	246	HEFK-----LSQASSKNNLLMARVSEITSPVLFREPPAASDTEHLIAIEILISTENNE	300
QY	238	OKSPRAIASSSYMPKHSIQARDHTGSMQHCFAQRCSQAASHSPMLPYENTNSEKPDPT	297
Db	301	ROPAPAL-----PRKRP	312
QY	298	KP-----DEKDVQNEWYIGEYSRQAVEDVLMKENKDGTFVLVRCSTKSKAEPYVL	348
Db	313	KPTVAANNSMNNMNSLODAEWYMGDISREVEYKEL-RDTADGTFVLVRASTKMGHD-YTL	370
QY	349	VVFYGNKYVYVYKIRPLESNOQFALGTGLRGNMPSVDIIENY	392
Db	371	TLRKGN--NKLKIFHRDGKYGFSDDL---TFNSVELLNHY	408
RESULT 11			
US-10-186-399-2			
Sequence 2, Application US/10186399			
Patent No. US20020173481A1			
GENERAL INFORMATION:			
APPLICANT: Ekman, Niklas			
APPLICANT: Arighi, Elena			
APPLICANT: Vastrik, Imre			
APPLICANT: Tamagnone, Luca			
APPLICANT: Alitalo, Kari			
TITLE OF INVENTION: REGULATION OF VASCULAR ENDOTHELIUM USING BMX TYROSINE			
TITLE OF INVENTION: KINASE			
FILE REFERENCE: 28113/31941A			
CURRENT APPLICATION NUMBER: US/10/186,399			
CURRENT FILING DATE: 2002-07-01			
PRIOR APPLICATION NUMBER: US 08/320,432			
PRIOR FILING DATE: 1994-10-07			
NUMBER OF SEQ ID NOS: 5			
SOFTWARE: PatentIn Ver. 2.0.			
SEQ ID NO 2			
LENGTH: 651			
TYPE: PRT			
ORGANISM: Mus musculus			
US-10-186-399-2			
Query Match 5.5%; Score 128; DB 9; Length 651;			
Best Local Similarity 24.1%; Pred. No. 0.28;			
Matches 61; Conservative 40; Mismatches 90; Indels 62; Gaps 11;			
QY	166	PAITLKKYQLPAPPESSAYFAFKPTPE-YONGPR-----ORSAKDFSRVL-----	214
Db	141	PGCTLWEAVADLHIALSDENKHR---APTFPEBLKIPRAVPLVKKMDASSGAILPQYDS	196
QY	215	-----GAEEESHQOTPRESSCPSSNQ--NTQKSPRAIASSTYMPKHSIQARDHTGSMQH	267
Db	197	YSKSCSGOPTSNIRIPIPEDCPDMWQVRLKLEEDIA-----	234
QY	268	CPAORCOAAASHSPRLPY---ENTNSEKPDPTKPEKDVQNEWYIGEYSRQAVEDVLM	324
Db	235	CSNQLERNTAISHTSKMSGCFRESSSE-----EENLHAYDWMFAGNITSQSQOLLR	287
QY	325	KENKDGTFVLVRCSTKSKAEPYLVF---YGNKYVYVYKIRPLESNOQFALGTGLGNEM	381

Db 288 QKQKGAEMVRN---SQMGMTVSLFSKAVNDKKGTVKHHVHTNAENKL--YLAENYC 342
QY 382 FDSVEDIEIHYTY 394
Db 343 FDSIKKLHHYH 355

RESULT 12

US-09-962-929-4
; Sequence 4, Application US/09962929
; Patent No. US20020115058A1
; GENERAL INFORMATION:

; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; APPLICANT: Nielsen, Anne A.
; TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated with A
; TITLE OF INVENTION: Expression of Plk3r1
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/668,644
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-962-929-4

Query Match 5.5%; Score 128; DB 10; Length 724;
Best Local Similarity 20.6%; Pred. No. 0.32;
Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;

QY 81 PSMKILPAPPIQSEYADTRFYQDMMEAPLLPPKAS----- 117
Db 85 PTPKPRPPRLPVAP-GSSKTEADVEQALTPDLAEOAPDIPAPLLIKLVEAIKKG 143
QY 118 -----VSTERQTRDVMTQLEVDKPTFK-----DVRSORFKGYTKINKTPPLPPR 165
Db 144 LECSTLYRTQSSNLAEIRQLDCTPSVDEMTDVHLADAFKRYL-----LPLPN 195
QY 166 PATLPKRYQPLPAPPESSATFAP-----KTFPEVQGRQRSARDEFVLGAEE 218
Db 196 PVIIPAAYSEIISLAPEYOSSEXYIQLKRLIRSPISIPHQYWLQYLKHFEXL----- 250
QY 219 ESHHQTPRESSCPSSNONTQKSPPAIASSYMPGKHSIQARDHTGSM----- 265
Db 251 -----SQTSSKNLMAARVLSEIFSPMLFRFSASSDNTENLIKVIILISTEWN 299
QY 266 --OHCPAQRQAAASHSPMLPYENTNSEKPDPTK-----DEKDVQWQNEWYIGY 314
Db 300 ERQPAPA-----LP-----PKPKPTTVANNNGNMNMSIQNMWYIGDI 358
QY 315 SROAVEVLAKENKDGTFVLVDCSTKSKAEPLYLVVFGKVVNVKTRFLESNOQFALGT 374
Db 339 SREEVNKL-RDTADGFLVLDASTKMHGD-YTLTKKGN-NKLKIFHRDQKYGFS 394
QY 375 GLRGNEFMDSVEDIEIHY 392
Db 395 PL-----TFSSVELINHY 408

RESULT 13

US-09-962-929-2
; Sequence 2, Application US/09962929
; Patent No. US20020115058A1
; GENERAL INFORMATION:

; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; APPLICANT: Nielsen, Anne A.
; TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated with A
; TITLE OF INVENTION: Expression of Plk3r1

FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/668,644
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-962-929-2

Query Match 5.5%; Score 127; DB 10; Length 724;
Best Local Similarity 22.0%; Pred. No. 0.38;
Matches 89; Conservative 45; Mismatches 148; Indels 122; Gaps 19;

QY 38 GRCAVLEPLPDHR--FNLGVPGEKCNSSNVDYD--PEFOLKAMPKMLPAPPIQES 94
Db 78 GRKR-ISPPTKPRPPRLPVAPSSKTEADTEQALPLDLAEOAPDIPAPLLIKL 136
QY 95 EYADTRFYQDMMEAPLLPPKASVSTERQTR-----DVMTQLEVDKPTFKDVRSORFK 149
Db 137 EAIKK-----GLECSTLYRTQSS-SNPABELRLDCAASVDLEMTDVHLADA---FK 187
QY 150 GFYTKINKTPLP-----PRAITLPKRYQPLPAPPE-ESSATFAPK 192
Db 188 --RYLADLPNFPVAVVANNEMSLAQELQSPEDCIQLKRLIRLPNIPHCWMLQYLLK 245
QY 193 PTFPEVQGRQRSARQD-----FSRYL-----GAEEESHQTKPRESSCPSSNQNT 237
Db 246 HFFK-----LSQASSKLNLAARVLEIFSPVLPFPAPASSDNTENLIKALIELISTEWN 300
QY 238 QKSPPAIASSSYMPGKHSIQARDHTGSMOHCPAQRQAAASHSPMLPYENTNSEKPDPT 297
Db 301 ROPAPAL-----PKPK----- 312
QY 298 KP-----DEKDVQWQNEWYIGYVROAVEVLMKENKDGTFVLVDCSTKSKAEPLY 348
Db 313 KPTTVANNSSNMNMSLODAEMWGDISREEVNKL-RDTADGFLVLDASTKMHGD-YTL 370
QY 349 VFEYGNKVNVKTRFLESNOQFALGTGLRGNEFMDSVEDIEIHY 392
Db 371 TPKKGN-NKLKIFHRDQKYGFSOPL-----TFNSVELINHY 408

RESULT 14

US-09-920-021A-3
; Sequence 3, Application US/09920021A
; Patent No. US20020110800A1
; GENERAL INFORMATION:

; APPLICANT: KAPLAN, DAVID
; APPLICANT: MARSH, H. NICK
; TITLE OF INVENTION: USE OF SHP-1 AND SHP-2 TO DETECT
; TITLE OF INVENTION: COMPOUNDS INVOLVED IN NEURONAL SURVIVAL
; FILE REFERENCE: 08338/016001
; CURRENT APPLICATION NUMBER: US/09/920,021A
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: US/08/918,157
; PRIOR FILING DATE: 1997-08-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-920-021A-3

Query Match 5.4%; Score 124.5; DB 10; Length 593;
Best Local Similarity 29.3%; Pred. No. 0.44;
Matches 43; Conservative 21; Mismatches 50; Indels 33; Gaps 5;

QY 295 DPTKPKDEKDVQWQNEWYIGYVROAVEVLMKENKDGTFVLVDCSTKSKAEPLYLVVFG- 353

```
Db 106 DPT-----SERWFHGLHSGKEAKLLETKGKHGSLVRE--SQSHGDFVLSVRTGD 155
QY 354 -----NKVYNYKIRFLESNOQFALGTGLRGNEMEDSVEDITEHYTFPIILIDGKD 404
Db 156 DKGESNDKSKVTHWIRCOE-----LKVDVGGERFDSLTDLVEHYKKNPMVETLG-- 207
QY 405 KAARRKQCYLTOPLPLARLLLTQYSSQ 431
Db 208 -----TVLQLKQPLNTTRINAAEIESR 229
```

RESULT 15

```
US-10-038-010-22
: Sequence 22, Application US/10038010
: Publication No. US20030040089A1
: GENERAL INFORMATION:
: APPLICANT: HYBRIGENICS
: APPLICANT: Pierre, Legrain
: TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
: FILE REFERENCE: BA767A
: CURRENT APPLICATION NUMBER: US/10/038,010
: PRIOR FILING DATE: 2002-07-23
: PRIOR APPLICATION NUMBER: US 60/259,377
: NUMBER OF SEQ ID NOS: 67
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 22
: LENGTH: 597
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: hSHP2_FL
: LOCATION: (1)-(597)
: OTHER INFORMATION:
US-10-038-010-22
```

Query Match 5.4%; Score 124.5; DB 9; Length 597;

Best Local Similarity 29.3%; Pred. No. 0.45; Mismatches 50; Indels 33; Gaps 5;

```
QY 295 DPTKDEKDVQNEWYIGEYSRQAVEDYLMKENKDGTEFLVRDCSTKSKAEPYLVVYFG- 353
Db 106 DPT-----SERWFHGLHSGKEAKLLETKGKHGSLVRE--SQSHGDFVLSVRTGD 155
QY 354 -----NKVYNYKIRFLESNOQFALGTGLRGNEMEDSVEDITEHYTFPIILIDGKD 404
Db 156 DKGESNDKSKVTHWIRCOE-----LKVDVGGERFDSLTDLVEHYKKNPMVETLG-- 207
QY 405 KAARRKQCYLTOPLPLARLLLTQYSSQ 431
Db 208 -----TVLQLKQPLNTTRINAAEIESR 229
```

Search completed: April 21, 2003, 12:39:41
Job time : 18.5549 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:33:12 ; Search time 16.6276 seconds
(without alignments)
2515.001 Million cell updates/sec

Title: US-09-856-061-2

Perfect score: 2316

Sequence: 1 MTSQGNKRTKEGFGDLRPO.....QPLPLALLTQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412.5	17.8	533	2 B56110	tyrosine phosphop
2	398	17.2	533	2 A56110	tyrosine phosphop
3	154.5	6.7	663	1 TVMYR	protein-tyrosine k
4	146	6.3	728	2 H59435	3-phosphatidylino
5	142.5	6.2	723	2 B38749	3-phosphatidylino
6	138.5	6.0	443	2 T27877	hypothetical prote
7	134.5	5.8	480	2 JC7552	Shb-like adapter p
8	133	5.8	1599	2 T15854	hypothetical prote
9	133	5.7	675	2 S60512	protein-tyrosine k
10	132.5	5.7	592	1 LBBY	actin-binding prot
11	132	5.7	1503	2 S17983	gene posterior sex
12	130.5	5.6	724	2 A38749	3-phosphatidylino
13	130	5.6	559	2 I49444	SH3 binding protei
14	130	5.6	724	2 A38747	phosphatidylinosi
15	129	5.6	751	2 S68957	adhesive plaque pr
16	128.5	5.5	3421	1 WZBBB6	367K tegument prot
17	128	5.5	706	2 A45990	junctional sarcopl
18	128	5.5	724	2 A38748	3-phosphatidylino
19	127	5.5	1385	2 T21706	hypothetical prote
20	125.5	5.4	1029	2 T30351	muclin-like protein
21	125	5.4	461	2 A57463	p55Epk - mouse
22	125	5.4	962	2 S58107	hypothetical WD-40
23	124.5	5.4	473	2 S25776	transforming prote
24	124.5	5.4	585	2 A46209	protein-tyrosine-p
25	124.5	5.4	593	1 JN0805	protein-tyrosine-p
26	124.5	5.4	597	1 A53593	protein-tyrosine-p
27	123.5	5.3	841	2 A43254	protein-tyrosine-p
28	123.5	5.3	1317	2 T03748	apoptosis associat
29	122.5	5.3	596	2 I38228	Shb protein - huma

30	122.5	5.3	875	2 S23760	polyphenolic adhes
31	121	5.2	485	2 T22147	hypothetical prote
32	120.5	5.2	431	2 T01557	hypothetical prote
33	120.5	5.2	1230	2 A36466	1-phosphatidylino
34	120.5	5.2	1792	2 A57075	tenin - chicken (
35	120	5.2	559	2 B44265	ENL (translocation
36	120	5.2	509	1 F0WGM	gag-abl polyprotel
37	119.5	5.2	595	1 A55651	protein-tyrosine k
38	119.5	5.2	1052	2 A82127	protein-tyrosine-p
39	119.5	5.2	1733	2 S27939	ribonuclease E VC2
40	119.5	5.2	1744	2 A54970	tenin - chicken
41	119.5	5.1	720	2 T26819	tenin, cardiac mu
42	119	5.1	2187	2 T30826	hypothetical prote
43	119	5.1	440	2 JC7807	nascent polypeptid
44	118.5	5.1	440	2 JC7807	Wiskott-Aldrich sy
45	118.5	5.1	657	2 A39526	probable protein k

ALIGNMENTS

RESULT 1	B56110	tyrosine phosphoprotein SLP-76 - mouse
C/Species:	Mus musculus (house mouse)	
C/Date:	28-Apr-1995	#sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C/Accession:	B56110	
R/Jackman, J.K.; Mottlo, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretz, J. Biol. Chem. 270, 7029-7032, 1995		
A>Title:	Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with	
A:Reference number:	A56110; PMID:95221345; PMID:7706237	
A:Accession:	B56110	
A>Status:	preliminary	
A:Molecule type:	mRNA	
A:Residues:	1-533 <JAC>	
A:Cross-references:	GB:U20159; NID:9806767; PIDN:AAC52189.1; PID:9806768	
C:Superfamily:	SH2 homology	
C:Keywords:	phosphoprotein	
F:422-520/Domain:	SH2 homology <SH2>	
Query Match	17.8%; Score 412.5; DB 2; Length 533;	
Best Local Similarity	29.7%; Pred. No. 2.1e-20;	
Matches	148; Conservative 51; Mismatches 175; Indels 125; Gaps 20;	
QY	14 FGLRLQNVSL-----KNRSMPSLSAKRCRAVLPLPDHRRNLG-----	56
DB	57 FPLRLMPLSLKSLQDINKNEERSIFTRKQIPRFLETFESHEDDGMSFEDDYESP	116
QY	57 --VPGGKCNNDYEDP--EFQL-----KAMP-----MKILPARPIQ--ESEY	96
DB	117 DDDPDGE---DDGDYSENEEQALVDAAADYEPSPNNEEALQSSITLPPNSFHTNSM	173
QY	97 ADTRYQDMMEAPLLP--PKASY-----	118
DB	174 IDRPPTKVSQPPVPLPRKPALPLPTGRNHSPLSPPHNHEBPSRSGNNTAKLPAP	233
QY	119 STERQTRDVMTOLEVD-----KPTKDYRSQRFKFTTKINKPLPPPPRAITL	170
DB	234 SIDRSTKPPDLRSIAPDRPFLTGKKPPSDPSAPLGRHLPKIQKPLP--PAMDR	290
QY	171 PKRYQPLPAPPRESSAYFAKPTPEVORGPRORSAKDPSRVLGAEESHHQTKRESSC	230
DB	291 HENNERLGPVTR-----KPSVPRHGGPDRRE-----NDDVDVHQRPPLQPSL	334
QY	231 PSSNONTQKSPPAIAS-----SYMPGKHSIQARDHTGSMQC--PAORCOAAASHSP	281
DB	335 PSMSSMTFFPSRSYQPSKNTFPLAHMGARS--ESNIGQGSASLPYFSGQPGNRPL	391
QY	282 -----RMLEPENTSEKPDPPKPEKDY--WONEYIGYISQAQVEDVLMKENKDGTLV	336
DB	392 RSEGRNLPLEVPN--RPQSPSPGEETPLDEEYVSATIRPEAEALRLKINQDGTFLV	449
QY	337 CSFKSAEPYVLVVFQGNKYVNVKIRFLESNQGFALGTGLRGNEMDSVEDIIEHTYFP	396

Db 450 SSKTANNPYVLMLVKOKVNIQIRYQESQVYLGTGLGKDEFLSVSDIIDYFRKMP 509
 Oy 397 ILLIDGKDKAARRKOCYIT 415
 Db 510 LLLIDGKNRGS-RYOCUTLT 527

RESULT 2

A56110
 Tyrosine phosphoprotein SLP-76 - human
 C:Species: Homo sapiens (man)
 C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
 R:Accession: A56110
 J:Biochem. 270, 7029-7032, 1995
 A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with G
 A:Reference number: A56110; MUID:95221345; PMID:7706237
 A:Accession: A56110
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-533 <JNC>
 A:Cross-references: GB:020158; NID:9806765; PIDN:AAC50135.1; PID:9806766
 C:Genetics:
 A:Gene: GDB:ICP2; SLP-76
 A:Cross-references: GDB:1230199; OMIM:601603
 A:Map position: 5g33.1-5qter
 C:Superfamily: SH2 homology
 C:Keywords: phosphoprotein
 F:422-520/Domain: SH2 homology <SH2>

Query Match 17.2%; Score 398; DB 2; Length 533;
 Best Local Similarity 30.2%; Pred. No. 2e-19;
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;

Oy 68 DYEDPEFQILKAMPKILPAPRIQES--YADTRYEDQMEAP-----LLPPKA 116
 Db 144 DYEPSPNDEEALQN-SILPAKFPNSNSMYIDRPPSGKTPOQPPVPQORPMALPPPA 202
 Oy 117 -----SVSTEROTRDVNRTOLEVD-----PPT 139
 Db 203 GRNHSPLPPPTQTHNEPSSRNHKTAKLPASIDSTKRPPLDRSLAPDRPREFTLGKKRP 262
 Oy 140 FKD---VRSQRFKPKYKINKTLPPLPPRPAITLPPKYOPLPAPPESSAYFAKPTFP 196
 Db 263 FSDKPSIFAGRSIGLHLKIQKPLP-----TERHRSPLP-----GKKPPVP 308
 Oy 197 EVQRPGRORSADKFSRVLGAEESHQTKPE-----SSCQSSNQNTQKSPPAIASSSY 250
 Db 309 KKGWCPDRRE-----NDEDDVHQRPDPALPMSNTPFSPSTKPSMNPPLPSSHM 360
 Oy 251 PGKHSIARQDHTGSMQHCPCAPQCOAASHSP-----RMLPYENTSEKPPDPKPEKYPW 305
 Db 361 PGAFS-ESSNSFPQASASLPYFSGSPSNKPPRAECGRNPLPLPKKPR-PSPAEESSL 418
 Oy 306 QNEWYIGEYSROAVEDVLMEKNDGTFLVDCSTKSKAEPLYLVYFGKKNVNVIRPLE 365
 Db 419 NEWYVSTITREAEALARKINQDGTFLVROSSKTTNPYVLMLYDKVNIQIRYOK 478
 Oy 366 SNOQALGTGLGNGEMFDSVEDIIEHYTFPILLIDGDKAARRKOCYIT 415
 Db 479 ESQVYLTGLGKDEFLSVSDIIDYFRKMPLLIDGKNRGS-RYOCUTLT 527

RESULT 3

TWVRR
 Protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rashee
 C:Species: feline sarcoma virus
 A:Note: host Felis sp. (cat)
 C:Date: 27-Nov-1985 #sequence_revision 26-May-1995 #text_change 31-Mar-2000
 C:Accession: A00653; A03937
 R:Nanarro, G.; Robbins, K.C.; Reddy, E.P.
 Science 223, 63-66, 1984

A:Title: Gene product of v-fgr onc: hybrid protein containing a portion of actin and
 A:Reference number: A00653; MUID:84097512; PMID:6318314

A:Accession: A00653
 A:Molecule type: DNA
 A:Residues: 1-663 <NAH>
 A:Cross-references: GB:X00255; GB:K01487; NID:961542; PIDN:CAA25063.1; PID:961543
 A:Note: the authors translated the codon GAT for residue 14 as Glu
 C:Comment: This protein is synthesized as a gag-fgr polypeptide.
 C:Genetics:
 A:Gene: fgr
 C:Superfamily: feline sarcoma virus protein-tyrosine kinase fgr; protein kinase homol
 C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; p
 F:1-118/Region: gag polypeptide similarity
 F:141-268/Region: actin similarity
 F:285-382/Domain: SH2 homology <SH2>
 F:402-660/Domain: protein kinase homology <KIN>
 F:410-418/Region: protein kinase ATP-binding motif
 F:432/Active site: Lys #status predicted

Query Match 6.7%; Score 154.5; DB 1; Length 663;
 Best Local Similarity 23.4%; Pred. No. 0.007;
 Matches 71; Conservative 34; Mismatches 91; Indels 107; Gaps 15;

Oy 161 LPPRPATLPKRY-----QPLPAP-----EESAYF-----189
 Db 101 LPPRPPTSLQPHSOPARALCRPVCRRPLPLPTAMEEVALVINDSGMCKAG 160
 Oy 190 -----AKPTPEYVGRGRSA-----KFSRYLGAEESHQTKRESSCPSSNQNTQ 238
 Db 161 FAGDDAPRAVPFISVGRHOGVWYGQKD--SYVGDAQSKR-----202
 Oy 239 KSPPAIASSYWPCKHSIOAR--DHTGSMQHCPCAPQCOAASHSPRMLPYENTSEKPD 295
 Db 203 ---GILTKY-PIEHGIVTMNDQEKIMHHTFVELNVAAEEHVLL-----TEAVLN 251
 Oy 236 PTKPEK-----DYNQEWYIGEYSROAV-DVLMEKNDGTFLV 334
 Db 252 PRANREKMTQIMFETFNIPSNVAVPDSIOAEWYFGKIGRDAEROLLSGNAGARLV 311
 Oy 335 RDCSTKSKAEPLYLVV-----FYGNKYVNVKIRFESNQOAFGLTGLGNGEMFDSVEDII 389
 Db 312 RESETTKA--YSLSIQMDKARQDVKHYKIRKIDTC-----GYITTRAFQNSVQELV 364
 Oy 390 EHY 392
 Db 365 QHY 367

RESULT 4

h59435
 Phosphoinositide-3-kinase regulatory beta chain [imported] - human

C:Species: Homo sapiens (man)
 C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: H59435; A59436
 R:Volinia S; Patrachini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L
 Oncogene 7, 789-793, 1992
 A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinosit
 A:Reference number: H59435
 A:Accession: H59435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <VOL>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1
 R:Janssen, J.W.; Schleithoff, L.; Bartram, C.R.; Schulz, A.S.
 Oncogene 16, 1767-1772, 1998
 A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta sub
 A:Reference number: A59436; MUID:98241181; PMID:9582025
 A:Accession: A59436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <JAN>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1

Query Match 6.3%; Score 146; DB 2; Length 728;
 Best Local Similarity 22.9%; Pred. No. 0.03;
 Matches 87; Conservative 48; Mismatches 147; Indels 98; Gaps 18;

QY 46 PLDDHRRNLGAVGGEKCNNDYDEDEPQ---LTKAMPKMK-----ILPARIQIE 93
 DB 95 PLUAPRRDGAPEPGLTLPDLPEQFSPDVAAPLVLVKEITKTLGDSHYPPELPAPR 154
 QY 94 SEYA-----DTRYFQDMMEAPDL-LP-----PKASVSTEROTRDVMTQLEEVDPKPT 140
 DB 155 TDMKSLSDVQMDTALADGKSLFLALPAPLVPEASAEKRLRLREAGVGALPEPTL 214
 QY 141 KDVRSQRFEGKYTKINKTPPLPPR--PAL-TLPKKYQPL--PPAPPESSAYFAPKPT 194
 DB 215 PLHRALTLR-FLLOHLGRVARRAPALGATFGLLRLNAPPPSPPPGAGADGS 273
 QY 195 FPEVQGPQRSAKDSRVLGAEESHQKPPSSCPSSNQNTQKSPPALASSYMPGKH 254
 DB 274 EP-----SPDEPALVETKLOEHLDEQVAPPALPPKPPKAPPA----- 312
 QY 255 SIQARDHTGSMQHPARQCAAAASHSPRLPYENTNSEKDPPTKPEKDVQWQNEWTGEX 314
 DB 313 -----PYYL-----ANGSP-----PSLQDA---EWYMGDI 335
 QY 315 SRQAVEDVLKKNKDGTFLLVRCSTKSKAEPYLVVFGKRYNVKIRPLESNOQFALGT 374
 DB 336 SREVENKTL-RDTPDGFLLVROASSKIQGE-YTLTLRKGN--NKLKVFHRGHGTFSE 391
 QY 375 GLRGNEMPSVEDIIEHYT 394
 DB 392 PL-----TFCSVDLNNHYRH 407

RESULT 5

B38749
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85k chain B - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 16-Jul-1999
 C:Accession: B38749
 R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, A.
 Cell 65, 91-104, 1991
 A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine ki
 A:Reference number: A38749; MUID:91191567; PMID:1707345
 A:Accession: B38749
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-723 <OTS>
 A:Cross-references: GB:M61745; GB:M61746
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:325-420/Domain: SH2 homology <SH2A>
 F:617-706/Domain: SH2 homology <SH2>

Query Match 6.2%; Score 142.5; DB 2; Length 723;
 Best Local Similarity 25.4%; Pred. No. 0.051;
 Matches 61; Conservative 23; Mismatches 79; Indels 77; Gaps 11;

QY 162 PPPRPAI-TLPKKYQPL-----PPAPPESSAYFAPKPTFPEVQGPQRSAKDSRVL 214
 DB 233 PAFGPAPVRLGATFGLLRLNAPPPSPPPG---APDGTPE-----TPDFPALL 278
 QY 215 GAEESHQKPPSSCPSSNQNTQKSPPALASSYMPGKSIQARDHTGSMQHPARQ 274
 DB 279 VEKLLQHELEQVAPPALPPKPPKTPA----- 307
 QY 275 AAASHSPRLPYENTNSEKDPPTKPEKDVQWQNEWTGEXSRQAVEDVLKKNKDGTFLL 334
 DB 308 -----PTGLANGSP-----PSLQDA---EWYMGDISREVENKTL-RDTPDGFLLV 349
 QY 335 RDCSTSKAEPYLVVFGKRYNVKIRPLESNOQFALGTLGKRNEMPSVEDIIEHYT 394
 DB 350 RDASSKIQGE-YTLTLRKGN--NKLKVFHRGHGTFSEPL-----TFCSVDLNNHYRH 402

RESULT 6

hypothenical protein ZK470.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27877
 R:Minx, P.
 submitted to the EMBL Data Library, October 1995
 A:Description: The sequence of C. elegans cosmid ZK470.
 A:Reference number: Z20433
 A:Accession: T27877
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-443 <MIN>
 A:Cross-references: EMBL:U39651; PDB:AAA80397.1; CESP:ZK470.5
 A:Gene: CESP:ZK470.5
 A:Introns: 39/3; 83/3; 109/3; 214/3; 270/3; 332/3; 352/3; 416/2

Query Match 6.0%; Score 138.5; DB 2; Length 443;
 Best Local Similarity 28.6%; Pred. No. 0.051;
 Matches 46; Conservative 23; Mismatches 49; Indels 43; Gaps 8;

QY 277 ASHSPRLPYENTNSEKDPPTKPEKDVQWQNEWTGEXSRQAVEDVLKKNKDGTFLLV 336
 DB 280 ASHOD-FAPOYSGNGELPMEQF-----WTFGRISERADLL-LHGRGEFLVRD 328
 QY 337 CSTKSKAEPYLVVFGKRYNVKIRPLESNOQFAL-----GTGLRG 378
 DB 329 -SESNKLOKLV-----EKLMDTKVVQTVAKPGOLISMGIERNKRFKQVNVNDGLKIG 381
 QY 379 NEMFDSVEDIIEHYTIFLLIDGKKAARKQCYLTQPL 419
 DB 382 NRTFVDMALNNHYTSPF-----SSPEKLFGLPLP 415

RESULT 7

JC7552
 Shb-like adapter protein, Shf - human
 C:Species: Homo sapiens (man)
 C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
 C:Accession: JC7552
 R:Lindholm, C.K.; Frantz, J.D.; Shoelson, S.E.; Welsh, M.
 Biochem. Biophys. Res. Commun. 278, 537-543, 2000
 A:Title: Shf, a Shb-like adapter protein, is involved in PDGF-alpha-receptor regulati
 A:Reference number: JC7552; MUID: 20548990; PMID:11095946
 A:Accession: JC7552
 A:Molecule type: mRNA
 A:Residues: 1-480 <LIN>
 C:Comment: This protein, a novel adapter protein, has the roles in PDGF-receptor and
 C:Genetics:
 A:Gene: Shf
 C:Keywords: apoptosis

Query Match 5.8%; Score 134.5; DB 2; Length 480;
 Best Local Similarity 20.7%; Pred. No. 0.11;
 Matches 85; Conservative 55; Mismatches 138; Indels 133; Gaps 19;

QY 26 KNSMPSLSSAK-GRCRAVLEPLP-----DHRNLGAVPGG-EKCNSENNDY 69
 DB 135 KKSMMRSTTLRLIRLRLEPRRLALLEYADPVDQETGEGSAGASGAPKPEPDGX 194
 QY 70 EDPFOLLKAMPKMLPAPRIQSEYADTRYFQDMMEAPDL-----PPKASVSTERQ 123
 DB 195 MEP-----YEAQKMM-----AETRGSKETATQPLPYDTPYEPEDGAPPEGE 237
 QY 124 -----TRDVRKTO-----LEEVDKP-TFKDVRSQRFKFKTKINKTLPPLP-----RPA 167
 DB 238 GAMPRESRLPEDEPPEEYDQFPEWKKERISAAFAVDIKVLDLWPPPPVQOLDSSPS 297
 QY 168 I-----TLPKKYQPLPAPPESSAYFAPKPTFPEVQGPQRSAKDSRVLGAEESH 222

Db 238 LPDGDRLDISGPASLPPEPSLEDDSSAOF----- 324
 Oy 223 QTKPSCPPSSNQNTQ-KSPRALASSSYMPGKHSTQARDHTGSMQHCPRACQAAASHSP 281
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
 C:Accession: T15854
 R: Fulton, L.
 submitted to the EMBL Data Library, June 1995
 A: Description: The sequence of C. elegans cosmid C56C10.
 A: Reference number: 218417
 A: Accession: T15854
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-1599 <EUL>
 A: Cross-references: EMBL:U29488; NID:9868238; PID:9868250; PIDN:AAA68780.1; CSDP:C56C10.
 C: Geneticals:
 A: Experimental source: Strain Bristol N2
 A: Gene: C56C10.12
 A: Introns: 38/3; 59/2; 92/3; 178/3; 380/1; 424/1; 573/3; 599/2; 697/1; 738/3; 980
 C: Superfamily: Caenorhabditis elegans hypothetical protein C56C10.12

RESULT 8
 T15854
 hypothetical protein C56C10.12 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
 C:Accession: T15854
 R: Fulton, L.
 submitted to the EMBL Data Library, June 1995
 A: Description: The sequence of C. elegans cosmid C56C10.
 A: Reference number: 218417
 A: Accession: T15854
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-1599 <EUL>
 A: Cross-references: EMBL:U29488; NID:9868238; PID:9868250; PIDN:AAA68780.1; CSDP:C56C10.
 C: Geneticals:
 A: Experimental source: Strain Bristol N2
 A: Gene: C56C10.12
 A: Introns: 38/3; 59/2; 92/3; 178/3; 380/1; 424/1; 573/3; 599/2; 697/1; 738/3; 980
 C: Superfamily: Caenorhabditis elegans hypothetical protein C56C10.12

Query Match 5.8%; Score 134; DB 2; Length 1599;
 Best Local Similarity 23.7%; Pred. NO. 0.51;
 Matches 56; Conservative 34; Mismatches 82; Indels 64; Gaps 13;
 Oy 85 ILPAPDIESEVADRYQDDMEAP-----LLLP-----PKASVSTRQTRDVRM 129
 Db 5 VRPKRP-KHREKQSDYPRPLDRAIVNGLRPAASHTIEDLERSASPEEDQDIST 63
 Oy 130 TQ---LEEVDPKTRFDVRSQRFKGYTKINKTLPPLPPRAITLPKKYQPLP----- 179
 Db 64 TEDSLAKREASBP-LKDYRSSVR-----PAPP-PRVSGERAPPLPPSMILFP 111
 Oy 180 -----APPESSAYFAPK-----PRPEVQGRFQRSKDSFVRL---GAEEES 220
 Db 112 RSTSMVAESRKESTAVAPKRSVAVASYPAVBELAELPSYDALQHPQVPSINGCLQHS 171
 Oy 221 HHQT---KPESSCPSSNQNTQ-KSPRALAS-SSTMGKHSIARDHTGSMQHCPR 269
 Db 172 HSATAIPKTRFSAVERERVRREGAPPMYPSIKTYERNEHGLMTEMLVTFYHNP 227

RESULT 9
 S60612
 protein-tyrosine kinase (EC 2.7.1.112) Bmx - human
 C:Species: Homo sapiens (man)
 C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 18-Feb-2000
 C:Accession: S60612
 R: Tamagone, L.; Lahtinen, I.; Mustonen, T.; Vitanneva, K.; Francis, F.; Muscatelli, F.;
 Oncogene 9, 3683-3688, 1994
 A: Title: Bmx, a novel nonreceptor tyrosine kinase gene of the BTK/ITK/TEC/TXK family loc
 A: Reference number: S60612; MUID:95060827; PMID:7970727
 A: Accession: S60612
 A: Status: preliminary; nucleic acid sequence not shown
 A: Molecule type: mRNA
 A: Residues: 1-675 <TAM>
 A: Cross-references: EMBL:X83107; NID:9951234; PIDN:CAA58169.1; PID:9951235

C: Geneticals:
 A: Gene: Bmx
 C: Key words: protein-tyrosine kinase tec; pleckstrin repeat homology; protein kinase
 F:3-109/Domain: pleckstrin repeat homology <PLK>
 F:296-392/Domain: SH2 homology <SH2>
 F:415-673/Domain: protein kinase homology <KIN>
 F:423-431/Region: protein kinase ATP-binding motif

Query Match 5.7%; Score 133; DB 2; Length 675;
 Best Local Similarity 24.0%; Pred. NO. 0.2;
 Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;
 Oy 166 PAITLPKTKQPLPPAPPESSAYFAPKPEPC-VORGP-----ORSKDFSRVGA--- 216
 Db 141 PCCTLMENAVANLHTVNEKRV-----PTFPRLKIPAVVLMKMDAPSSSTTLAQTDN 196
 Oy 217 EESHQTKPESSCPSSNQNTQKSPRALAS-----SYMPG-----KHSIQAR 259
 Db 197 ESKKHYGSPSSSTSLAQYDSNKKIGSQPNFMQYIRPDPDMQVRLKSSSSSE 256
 Oy 260 DHTGSMQHCPRACQAAASHSPMLPYENTNSEKDPDKEDYQWQEWYIGYSRQAV 319
 Db 257 DVASSNQK-----BRNVNHTTSKISWEPSSSE---EENLDYDWFEGNISRSQS 306
 Oy 320 EDVLMKENKDGTEFLVRCSTKSKAPVYLVE---YGNKYVWKTRPLESNOQFALGTGL 376
 Db 307 EQLLNQKQKREGAFWVRN---SSQVGMVYVLSFKAVNDKKGVKHYHTNAENKL--YL 361
 Oy 377 KGENMPSDVEDIIEHYTY 394
 Db 362 AENVCFDSIPKLIHYHOH 379

RESULT 10
 LIBY
 actin-binding protein ABP1 - yeast (Saccharomyces cerevisiae)
 N: Alternate names: protein YCR088w
 C: Species: Saccharomyces cerevisiae
 C: Date: 31-Mar-1991 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
 C: Accession: S19503; S19767; S07608
 R: Dusterhoft, A.; Erdmann, D.; Hegemann, J.; Philippson, P.; Schweitzer, B.; Spiegelb
 submitted to the Protein Sequence Database, March 1992
 A: Reference number: S19351
 A: Accession: S19503
 A: Molecule type: DNA
 A: Residues: 1-476 <DUS>
 A: Cross-references: EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
 R: Frontali, L.; Grisanti, P.
 submitted to the Protein Sequence Database, March 1992
 A: Reference number: S19504
 A: Accession: S19767
 A: Molecule type: DNA
 A: Residues: 14-592 <PRO>
 A: Cross-references: EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
 R: Drubin, D.G.; Mulholland, J.; Zhu, Z.; Botstein, D.
 Nature 343, 288-290, 1990
 A: Title: Homology of a yeast actin-binding protein to signal transduction proteins an
 A: Reference number: S07608; MUID:90136906; PMID:2405279
 A: Accession: S07608
 A: Molecule type: DNA
 A: Residues: 1-57,'S',59-111,'I',313-592 <DRU>
 A: Cross-references: EMBL:X51780; NID:933321; PIDN:CAA36075.1; PID:933322
 C: Geneticals:
 A: Gene: SGD:ABP1; MIPS:YCR088w
 A: Cross-references: SGD:S0000684; MIPS:YCR088w
 A: Map position: 3R
 C: Superfamily: actin-binding protein ABP1; SH3 homology
 C: Keywords: actin binding; duplication
 F:158-180,378-401/Region: duplication
 F:339-388/Domain: SH3 homology <SH3>

Query Match 5.7%; Score 132.5; DB 1; Length 592;

Best Local Similarity	21.48;	Pred. No.	0.19;
Matches	54;	Conservative	44;
		Mismatches	99;
		Indels	55;
		Gaps	11.

```

QY 117 SVSREOTBDVBMQLEVEVDPTFKDVRSGKGFYTKIKPTL-----PPR---PAI 168
      | : : : | : : | : : : | : : | : : | : : | : : | : : | : : |
Db 329 NAEKAEKPKVEVEDEGEDEPDYDKLS-KFEGLASKEKEEMENKFAFPKPKSEPTI 387
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 169 TLPKKY-OPLPADPESSAYFAKPTFVEYQGRPRORSADKDFSVLGAEEESHQTKPE 227
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 388 ISPFKFSKPEPVKAEAE--QPTDYKKIG-----PLDPMHTEADNEEPE 433
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 228 SS-----QPSN-----ONTOKSPALASSVMPGHSIQARDHNGSM 265
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 434 ENDDMDDEDEEAQPLPSRNVA SGAPVQKEEPEQETIAS--LPSSNSIPAPKQEPAP 491
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 266 QHCEAQRQAASHSPMLPYENTNSEKPDPTK-PDEKDVQNMENWYIGEYSRQAEVDYLM 324
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 492 EQAPEEELTEEAEEBAAPOLPFRSSAAPPPRRATPEKKPKENPMWATAYDYDAED--- 548
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 325 KENKDGTFELVD 336
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 549 ---NELTFEVD 557
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : |

```

RESULT 11

gene posterior sex combs protein - fruit fly (Drosophila melanogaster) SI/983

N:Alternate names: regulatory protein Psc

C:Species: Drosophila melanogaster

C:Date: 22-Nov-1993 #sequence,revision 26-May-1995 #text-change 21-Jul-2000

C:Accession: SI7983; SI7880

R:Brunk, B.P.; Martin, E.C.; Adler, P.N.

Nature 353, 351-353, 1991

A:Title: Drosophila genes Posterior Sex Combs and Suppressor two of zeste encode proteins

A:Reference number: SI7983; MUID:92018190; PMID:1833647

A:Accession: SI7983

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1603

A:Cross-references: EMBL:X59275; NID:98357; PIND:CA441965.1; PID:98358

A:Note: translation of nucleotide sequence is not complete

R:van Lohuizen, M.; Frasch, M.; Wientjens, E.; Berns, A.

Nature 353, 353-355, 1991

A:Title: Sequence similarity between the mammalian bmi-1 proto-oncogene and the Drosophila

A:Reference number: SI7880; MUID:92018191; PMID:1922340

A:Accession: SI7880

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 231-452, 'I', 454-591 <LOH>

C:Genetics:

A:Gene: FlyBase:Psc

A:Cross-references: FlyBase:FBgn0005624

C:Superfamily: RING finger homology

C:Keywords: DNA binding; nucleus

F:261-309/Domain: RING finger homology <RNG>

Query Match	5.7%;	Score 132;	DB 2;	Length 1603;
Best Local Similarity	21.2%;	Pred. No. 0.7;		
Matches	80;	Conservative	51;	Mismatches 148;
			Indels	98;
			Gaps	18

```
QY      26 KNRSPSLSSAKGCR-----AVLEPLPDHRRNLAGVPG-----GECKNSNNDYEDEFOL 76
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      641 KSSSSSSSSSGEKKRKRKSPPLTPPLTIRTERIMSPSGVSTLSPRTSAGFSEDPKSEF 700
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      .77 LKAMP-----SMKILPARP-IOESEVADR--YFOO--WMAAPLLLPK 115
        || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      701 LKSFLLAKTKYKVESPERTLNRRATPPSPVSQOSAPKSNGNNLDLSILMKPPSCAPK 760
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      116 ASVTROTTRD----VRMTOLEVDKP--FK---DVRSOFKFGFYKTINKTPLPAR-R 165
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      761 SIASSRKSRKEVKVASKOKLSPPLPYDPKINLLPTNGNSSOSTASPKIEKLMLMPPAK 820
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      166 PAITLPKYQPL-----PPADPESSA--YFAP-----KTFPEVOR-GPROSA 207
```

```

Db      821 PRLALPKLQPSAQFAPRPSPIHHAGVQMSAAGNTPRIKKRYQPLPKASRRNPANIP 880
QY      208 KDFSRVL-----GAEEESHHQTPRESSCPSSNONTQKSPALASSYMPGRK 254
Db      881 NDVAKRLKDACTELKSTGGGSVEVNNSSAQKPHLYGKG--ETKMGPRALPATTPGQGN 938
QY      255 SIQARDHTGSMQCP-----AQRCAASHSPMLPYENTNSEKDDPTK 298
Db      939 NVGQAQGLNPLPSAPPKNGSSNNYLNALNLNSMKCKGEAPPCGCRPTMYTPNSPIYSS 998
QY      299 PDEKDVQWQNEWYIGEYS 315
Db      999 PQ-----YVPSYN 1006

```

RESULT 12
A30740

A36749
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain alpha - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 05-Nov-1999
 C:Accession: A38749
 R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson,
 Cell 65, 91-104, 1991
 A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine
 A:Reference number: A38749; MUID:91191567; PMID:1707345
 A:Accession: A38749
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-724 <OMS>
 A:Cross-references: GB:M61746; GB:M61745; NID:q163476; PIDN:AAA75911.1; PID:q163477
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 E:333-428/Domain: SH2 homology <SH2A>
 E:624-718/Domain: SH2 homology <SH2>

Query Match	5.6%	Score 130.5	DB 2	Length 724
Best Local Similarity	20.2%	Pred. No. 0.33		
Best Match 72	Conservative 43	Mismatches 114	Indels 127	Gaps 15

OY	148	FKGKTYKINTPLRPPRPATLTKKQOPLPAP-PESSA-----	YAPK	192
Db	69	FPGYUYEYIGKKKISSPPK---PRPPRLVVAAGPSKTEADSEQOASTLPDLAEQFAP		1253
OY	193	PTFPEVOGPRORSKADQESRYLGAEEESNOTKPESS-----	CPSSNONTQ	238
Db	126	DVAP-----PLLIKAVEIEKKGECSUTLYUOTSSSNAPBLROLDDCTFSLDEMDVH		180
OY	239	-----KSP--PALASSYMGKHSIARD-----HTGSMOCHQAPQC		2737
Db	181	VLADAFKRYLLDLPYUYPVAVSSSELLSLAEVOSSEYIOLKKLIRSPISIPQWYLT		2400
OY	274	QAAASH-----SPMLDYENTNSEK-----		2939
Db	241	QYLLNHFPEKLSQTSKKNLNAVLSSELLSPLLFFRPAASSENTEHLIKIIEILISTEMNE		3000
OY	294	-----PDBTKP-----DEKVOQWQEWYIGESROAVEDVLMKEKQDTFLVRD		3366
Db	301	ROPAPALPPKPPKPTTYUANNGNNMNMMSIODEKTYWGDLSREVAEKL-RDAPDSTFLVRD		3539
OY	337	CSTKSKAEPUYLVUYFGNKYUVAKIRFELSSMOQFALGTGLGEMFDSVEDIIEHY		392
Db	360	ASTKMHGQ-YLLTEKCGN--NKLIKIFGNQKQGFSDPL-----TFMSVVELLNHY		408

RESULT 1

149444
SH3 binding protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #next_change 05-Nov-1996
C:Accession: 149444
R:Ren, R.; Mayer, B.J.; Cicchetti, P.; Baltimore, D.

Science 259, 1157-1161, 1993
 A>Title: Identification of a ten-amino acid proline-rich SH3 binding site.
 A:Reference number: 149444; MUID:93174278; PMID:8438166
 A:Accession: 149444
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-559 <RES>
 A:Cross-references: GB:L14543; NID:9293267; PIDN:AAA37121.1; PID:9293268
 C:Superfamily: pleckstrin repeat homology
 F:25-128/Domain: pleckstrin repeat homology <PLK>
 F:201-210/Region: proline-rich SH3 binding

Query Match 5.6%; Score 130; DB 2; Length 559;
 Best Local Similarity 21.4%; Pred. No. 0.26;
 Matches 87; Conservative 42; Mismatches 145; Indels 132; Gaps 19;

OY 86 LPARIQSEY---ADRRYQDMMEAPLL-----PPKASSTEGTQDVNMT 130
 DB 164 LSSYFMNEDYHEDEDDSYLEPSPGPKLEDAITYPAPYPPVPR----- 213
 OY 131 QLEVDKPTFKDVSQRFKGYTKINKTP-LPPRPATITP-----KKY 174
 DB 214 -----KRAFSDL-----RAHSFTSKSPSPLPPPPKRGLPDTSAPEDAKDALGLKRV 263
 OY 175 QP---LPPAPPESSAIFYAPKFTPEYQGRORS-----A 207
 DB 264 EPLGLVPATPRRMSPPMNPTVPLNLRKPCFSDSNVNGLEPMTPGHGTSSSSTMA 323
 OY 208 KDFSVLGAEEESHQTK--PESSCPSSNQNTQ-----SPRAISSSTMPCKHSIQ 257
 DB 324 VATSNCKOKLKFLLSSGPPSEPPVPANPKFLKAEPSPREAKFAVP---PVA 380
 OY 238 ARDHTGSMQHCPCQAASHSPMLPYENTNSEKPD-----PTKPD--- 300
 DB 381 PRPPQKMP-MPEATVRAVLPREPNTPLPHLQSPRPGQSRGFSFKANQPSQADGE 439
 OY 301 -----EKDWQNMWTGYEYSRQAVEDVLMK-----ENKQGTFLVRCSTSKSAEPYVL 348
 DB 440 EDSDEYKRVPLPNSVFNTTESCVERLFRATDPRGEPQGLCYCIRNSFTKSGK---VL 496
 OY 349 VFE---YGNKYVNVKIRFLESNOQFALGTLGNEMEDVEDIIEHY 392
 DB 497 VVWDESSNKVKRYRI--FEKOSKF-----LGEVLFASVGSNVEY 536

RESULT 14
 A38747
 C:phosphatidylinositol 3-kinase (EC 2.7.1.-) 85k chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 16-Jul-1999
 C:Accession: A38747
 R:Escobedo, J.A.; Navankasattusas, S.; Kavanaugh, W.M.; Milfay, D.; Fried, V.A.; William
 Cell 65, 75-82, 1991
 A:Title: cDNA cloning of a novel 85 kd protein that has SH2 domains and regulates bindin
 A:Reference number: A38747; MUID:91191564; PMID:1849460
 A:Accession: A38747
 A:Molecule type: mRNA
 A:Residues: 1-724 <ESC>
 A:Cross-references: GB:M60651
 C:Comment: This protein binds a phosphotyrosine-containing sequence of ligand-activated
 C:phosphatidylinositol at position 3 of the inositol ring.
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:33-428/Domain: SH2 homology <SH2A>
 F:624-718/Domain: SH2 homology <SH2>

Query Match 5.6%; Score 130; DB 2; Length 724;
 Best Local Similarity 22.0%; Pred. No. 0.35;
 Matches 89; Conservative 46; Mismatches 147; Indels 122; Gaps 19;

OY 38 GRCAVLEPLDHR--RNLGVGCGEKNSNND--PEFQILKAWPSMKILPARIQES 94
 DB 78 GRKR-ISPPTKPRPPRLPVAPOSSSKTEADTEQALPLDLAQFAPDPVAPLLIKL 136

OY 95 EXADTRYQDMMEAPLLPPKASSTEGTQ-----DYRMQLEEDVKPTFKDVSQRFK 149
 DB 137 EAIKK-----GLEGTSLVTRTQSS--SNPAELRLDCCDAASVLEMDIVLADA---FK 187
 OY 150 GFKYTKINKTPLP-----PPRAITLPKKYQPLPPAPPE-ESSAYFAPK 192
 DB 188 --RYLADLPNPVITPVAAYVNEKMSLAQELQSPEDCQLLKLRLPNIRHQCLTLQIYLK 245
 OY 193 PFPEYQGRORSKAD-----FSRVL-----GAEEESHQTPRESSCPSSNONT 237
 DB 246 HFK-----LSQASSKNILNARVLSEIFSPVLFPPAASDSNTEHLIKAILISTENNE 300
 OY 238 QKSPRAISSSTMPCKHSIQARDHTGSMQHCPCQAASHSPMLPYENTNSEKPDPT 297
 DB 301 KQAPAL-----PPKPP 312
 OY 298 KP-----EKDWQNMWTGYEYSRQAVEDVLMKENDGTFLVRCSTSKSAEPYVL 348
 DB 313 KPTVANNMNNMNSLQDAEWGDISREVENKL--RDTAGTFLVRASTKMHGD-YTL 370
 OY 349 VVFYGNKYVNVKIRFLESNOQFALGTLGNEMEDVEDIIEHY 392
 DB 371 TLKRGGN--NKLKIKIFHRDGKGFSDPL-----TFNSVELINHY 408

RESULT 15
 S68957
 C:adhesive plaque protein precursor - Mediterranean mussel
 C:Species: Mytilus galloprovincialis (Mediterranean mussel)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 18-Feb-2000
 C:Accession: S68957
 R:Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.; Haraizame, S.; Waite, J.H.
 submitted to the EMBL Data Library, August 1995
 A:Description: Cloning, sequencing and sites of expression of genes for the hydroxyl
 A:Reference number: S68957
 A:Accession: S68957
 A:Molecule type: mRNA
 A:Residues: 1-751 <INOC>
 A:Cross-references: EMBL:D63778; NID:9961463; PIDN:BA00981.1; PID:d1010497; PID:g9661

C:Keywords: hydroxyproline
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-751/Product: adhesive plaque protein #status predicted <MAT>
 F:123,127,133,173,177,183,227,267,283,293,297,303,307,313,323,327,333,337,343,347,367
 F:33/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:125,129,135,139,175,179,185,189,225,229,265,269,285,289,295,299,305,309,315,319,325
 F:79/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status predicted
 F:126,136,176,186,226,266,286,296,306,316,326,336,346,366,376,386,406,416,426,436,446
 F:ans-2,3-cla-3,4-dihydroxyproline (Pro) #status predicted
 F:485,489,495,499,505,509,515,519,525,529,535,539,545,549,555,559,565,569,575,579,585
 F:485,489,495,499,505,509,515,519,525,529,535,539,545,549,555,559,565,569,575,579,585
 F:537,543,547,553,567,573,577,583,587,593,597,603,607,613,617,623,627,633,637,643,647

Query Match 5.6%; Score 129; DB 2; Length 751;
 Best Local Similarity 23.4%; Pred. No. 0.43;
 Matches 55; Conservative 34; Mismatches 108; Indels 38; Gaps 9;

OY 84 KILARPQSEYVADTRY---FQDMMEAPLLPPKASSTEGTQRCQRMQDLEVDKPTF 140
 DB 88 KLSSTYKPLKTYNKNKTNPVYKPKMTYPTTKKPSPTPYK-----PKPSY 135
 OY 141 KDVSORFKGFKYTKINKTP-LPPRPATITLPKKYQPLPPAPPESSAYFAPKPTPEYQOR 200
 DB 136 PATYKSSSYPSYKPKKTYPTTKPLTLTPPYTKKPPSYPT-----YKPKPYPATYK 190
 OY 201 G-----PRQSAKDFSRVLGAEEESHQTKPESSCPSSNONTQKSP-----AIASSSYM 250
 DB 191 SKSSYPSYKPKKTYPSYKPKTYPTSTYKPKVSYPTPYKSKKSYPTIYTKAKYSPSYK 250
 OY 251 PGKHSIQARDHTGSMQHCPCQAASHSPMLP-YENTNSEKPD--PT---KP 299
 DB 251 PKK--TYPSTYKPKISTYPTTKAKPSYPTSTYKAKPSYPTTKAKPSYPTTKAKP 303

Tue Apr 22 16:29:16 2003

us-09-856-061-2.rpr

Page 7

Search completed: April 21, 2003, 12:38:34
Job time : 20.6276 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:28:57 ; Search time 9.65475 Seconds

(without alignments)
1868.739 Million cell updates/sec

Title: US-09-856-061-2

Sequence: 1 MTSQGNKRTKEGFGDIRFO.....QPLPLARLLITQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412.5	17.8	533	1	LCP2_MOUSE
2	398	17.2	533	1	LCP2_MOUSE
3	146	6.3	724	1	P85B_BOVIN
4	146	6.3	724	1	P85B_BOVIN
5	143.5	6.2	722	1	P85B_RAT
6	142	6.1	545	1	FCR_FSVGR
7	140	6.0	540	1	SCR_HUMAN
8	136.5	5.9	561	1	BMX_HUMAN
9	133	5.7	675	1	P85A_RAT
10	133	5.7	724	1	P85A_RAT
11	132.5	5.7	217	1	GRAP_HUMAN
12	132.5	5.7	271	1	CRK_DROME
13	132.5	5.7	592	1	ABP1_YEAST
14	132	5.7	1603	1	PSC_DROME
15	130.5	5.6	722	1	P85B_MOUSE
16	130.5	5.6	724	1	P85A_BOVIN
17	130	5.6	559	1	3BP2_MOUSE
18	130	5.6	724	1	P85A_MOUSE
19	129	5.6	461	1	P85G_HUMAN
20	129	5.6	751	1	FPL_MYTG
21	128.5	5.5	3421	1	TEGU_HSVB
22	128	5.5	705	1	TRDN_RABIT
23	128	5.5	724	1	P85A_HUMAN
24	127	5.5	461	1	P55G_BOVIN
25	127	5.5	1362	1	BRD4_HUMAN
26	125	5.4	461	1	P55G_MOUSE
27	125	5.4	962	1	YADE_SCHPO
28	124.5	5.4	583	1	SHC_HUMAN
29	124.5	5.4	585	1	PTNB_MOUSE
30	124.5	5.4	593	1	PTNB_HUMAN
31	123.5	5.3	845	1	CSW_DROME
32	123.5	5.3	1386	1	ZAP3_MOUSE
33	123.5	5.3	1386	1	ZAP3_MOUSE

34	122.5	5.3	875	1	FPL_MYTED	025460 mytilus edu
35	120.5	5.2	1290	1	PIGL_HUMAN	P19174 homo sapien
36	120	5.2	559	1	ENL_HUMAN	003111 homo sapien
37	119.5	5.2	509	1	STK_HYDAT	P17713 hydra atten
38	119.5	5.2	1490	1	CRK7_HUMAN	09nyv4 homo sapien
39	119.5	5.2	1490	1	TENS_CHICK	004205 gallus gall
40	119	5.1	1744	1	FGR_MOUSE	P14234 mus musculu
41	118.5	5.1	526	1	SRC_RSVSR	P00524 rous sarcom
42	118.5	5.1	657	1	SGV1_YEAST	P23293 saccharomyc
43	118.5	5.1	659	1	BTK_HUMAN	006187 homo sapien
44	118	5.1	2774	1	MAPA_RAT	P34926 rattus norv
45	117.5	5.1	659	1	BTK_MOUSE	P35991 mus musculu

ALIGNMENTS

RESULT 1
LCP2_MOUSE
ID LCP2_MOUSE STANDARD; PRT; 533 AA.
AC Q60787;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell lymphoma;
RX MEDLINE=95221345; PubMed=7706237;
RA Jackman J.K., Motta D.G., Sun Q., Tanemoto M., Turk C.W., Peltz G.A., Koretzky G.A., Finkel P.R.;
RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells";
RL J. Biol. Chem. 270:7029-7032(1995).
CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND Fyb.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES.
CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: U20159; AAC52189.1; -
CC HSSP: P12931; ISHD.
CC MGD: MGI:1321402; LCP2.
CC InterPro: IPR001660; SAM.
CC InterPro: IPR00980; SH2.
CC Pfam: PF00017; SH2; 1.
CC Prodom: PD000093; SH2; 1.
CC SMART: SM00454; SAM; 1.
CC SMART: SM00252; SH2; 1.
CC PROSITE: PS00001; SH2; 1.
CC SH2 domain; Phosphorylation.
CC DOMAIN 422 530
CC SEQUENCE 533 AA; 60228 MW; 60DB77782C2E60C3 CRC64;
Query Match 17.8%; Score 412.5; DB 1; Length 533;
Best Local Similarity 29.7%; Pred. No. 1.7e-19;

Matches 148; Conservative 51; Mismatches 175; Indels 125; Gaps 20;

QY 14 FGDRLFQVNSL-----KNRSMPSLSAKGRCRAVLPLPHRRNLG-----56
 DB 57 FPKLMLPLSLSDINKNEERRSTFTTKPPIRPLETETSHEDDGGWSSFEEDYESPN 116
 QY 57 --VEGEGCNSNDYEDP--EFOLI-----KAMPS-----MKILPAPPIQ--ESEY 96
 DB 117 DDDPGE---DDGYESPNEEEOALVDDADYEPSPNNEBALQSLIPSPFNNTSMY 173
 QY 97 ADTRFYQDMAPLPLP--PKASY-----118
 DB 174 IDREPTKVSQPPVPPPLRKPAPLPLPTGNNHSLPSLPPHNEEPSRGNNKTAKLPAP 233
 QY 119 STERQTRDVRMTQLEVD-----KTFEKDVSORFKGKRYKINKTPLPPRPAITL 170
 DB 234 SIDRSTKPLDRLSLAPLDREPLTGKKRPFSDKPSAPLGRHRLKIQKPLP---PAMDR 290
 QY 171 PKKYOPLPAPPESSAVFAKPFEPVQGPORRSKADSRVLGAEEESHQTPRESSC 230
 DB 291 HERNERLGPVTR-----KPSVPRHGRCPDRE-----NDEDDVHQRPPLQPSL 334
 QY 231 PSSNONTOKSPPAIASS-----SYMGRHSIQARDHTGSMOHC--PAORCOAAASHSP- 281
 DB 335 PSMSSNTFSPRSVQPSKNTFPLAHMGAFS---ESNIGQOASLPPYFSQCGGNPPL 391
 QY 282 ----RMLPYENTSEKPDPTKPEKDV--WONWYIGESQOAVEDVLMKENKDTFLVRD 336
 DB 392 RSEGRNLPPLVPN--RQPPSPGEEETPLDEEYVSYTRPEAAALRKINQOTFLVRD 449
 QY 337 CSTSKRPEPYLVVYFNGKNVYKIRFLESNOQFALGTGRGMEFVSVDIEHYTFP 336
 DB 450 SSKKTANPFLVWLVDKYNIOIRYQESOVYLLGTGLRGKEDFLSVSDIIDYFRKMP 509
 QY 397 TLIDGKDKARKKQCYLT 415
 DB 510 LLLIDGKNRGS-RYQCTLT 527

RESULT 2
 LCP2_HUMAN STANDARD; PRT; 533 AA.
 ID LCP2_HUMAN
 AC Q13094;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
 GN LCP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Leukemia;
 RX MEDLINE=95221345; PubMed=7706237;
 RA Jackman J.K., Motto D.G., Sun Q., Tanemoto M., Turk C.W., Peltz G.A., Kortsitz J.A., Finkel P.R.,
 RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells."
 RL J. Biol. Chem. 270:7029-7032(1995).
 RP SEQUENCE FROM N.A.
 RC TISSUE=prostate;
 RA Strusberg R.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
 CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND FIB.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES. HIGHLY EXPRESSED ALSO IN T CELL AND MONOCYTIC CELL LINES, EXPRESSED AT LOWER LEVEL IN B CELL LINES.

CC NOT DETECTED IN FIBROBLAST OR NEUROBLASTOMA CELL LINES.
 CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION BY ZAP-70.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- DATABASE: NAME=PRO; NOTE=PRO 1:1-5(2000);
 CC WWW=ftp://www.ncbi.nlm.nih.gov/prov/guide/1118450040.g.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-stb.ch).
 CC
 DR EMBL: U20158; AAC50135.1; -
 DR EMBL: BC016618; AAH16618.1; -
 DR HSP: P12931; ISHD.
 DR Genew: HGNC:6529; LCP2.
 DR MIM: 601603; -
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2.1.
 DR ProDom: PD000093; SH2.1.
 DR SMART: SM00454; SAM; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2.1.
 DR SH2 domain; Phosphorylation.
 FT DOMAIN 133 136 POLY-GLU.
 FT DOMAIN 198 201 POLY-PRO.
 FT DOMAIN 422 530 SH2.
 SQ SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;

Query Match 17.2%; Score 398; DB 1; Length 533;
 Best local similarity 30.2%; Pred. No. 1.5e-18;
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;

QY 68 DYEDPEFQLKAMPKMKILPAPRIQSE--YADTRYFQDMENP-----LLPDKA 116
 DB 144 DYEDPPSNDDEALQN-SILPAKFPNSNMYIDRPPSGKTPOQPPVPPQRPMAALPPPA 202
 QY 117 -----SYTEROTRYVMTQLEVD-----KPT 139
 DB 203 GRNHSLPPTQTNEEPSRSRNHKTAKLPAPSIDRSTKPLDRLSLAPFREPTLCKKP 262
 QY 140 FKD--VRSORFKGFYTKINKTPLPPRPAITLPPKYQPLPAPPESSAVFAKPTP 196
 DB 263 FSDKPEIPAGRSIGELPLRIQKPLP-----TTERHRSPLP-----GKKPPV 308
 QY 197 EYQGRGROASAKDFSRVLGAEEESHQTKPE-----SSCPSSNONTOKSPPAIASSSY 250
 DB 309 KKGWGPDRRE-----NDEDDVHQRPPLQPALPMSSTNTPSPRSSTKPSPNPLPSSIM 360
 QY 251 PGKHSIOARDHTGSMOHCPAORCOAAASHSP-----RMLPYNTSEKPDPTKPEKDVW 305
 DB 361 PPAFS-ESNSSPPQASLPPYFSQGPSNRPRAACRNPPLPAPKPPR-PSPAEENS 418
 QY 306 QNEWYIGESQOAVEDVLMKENKDTFLVDCSTSKAKAPYLVVYFNGKNVYKIRFLE 365
 DB 419 NEWYVSYTRPEAAALRKINQDGTFLVROSSKTTNPVLYWLVDKYNIOIRYQK 478
 QY 366 SNOQALGTGLGNGEMFVSVDIEHYTFPPLIDGDKARKKQCYLT 415
 DB 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLIDGKNRGS-RYQCTLT 527

RESULT 3
 P85B_BOVIN STANDARD; PRT; 724 AA.
 ID P85B_BOVIN
 AC P23726;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase

DE p85-beta subunit) (Ptdins-3-kinase p85-beta).
GN PIK3R2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91191567; PubMed=1707345;
RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Larrea F., Panayotou G.,
RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
RA Coutnidge S.A., Parker P.J., Waterfield M.D.;
RT "Characterization of two 85 kd proteins that associate with receptor
RT tyrosine kinases, middle-T/pp60-src complexes, and p13-kinase.";
RL Cell 65:91-104(1991).
RN [2]
RP CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.
RX MEDLINE=93049176; PubMed=1330535;
RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,
RA Fry M.J., Blundell T.L., Wolmer A., Waterfield M.D.;
RT Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
RT domain with a PDGF receptor phosphorylation site: structural features
RT and analysis of conformational changes.";
RL EMBO J. 11:4261-4272(1992).
CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M61746; AAA79510.1; -
CC HSSP: P23727; 2PNB.
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000198; RhogAP.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF000017; SH2; 2.
DR Pfam: PF000018; SH3; 1.
DR Pfam: PF00620; RhogAP; 1.
DR PRINTS: PR00678; PI3KINASEP85.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000066; SH3; 1.
DR PRODOM: PD000093; SH2; 2.
DR SMART: SM00324; RhogAP; 1.
DR SMART: SM00252; SH2; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50001; SH2; 2.
DR PROSITE: PS50002; SH3; 1.
KM SH3 domain; SH2 domain: Repeat.
FT DOMAIN 4 80
FT DOMAIN 125 255 RHO-GAP.
FT DOMAIN 326 421 SH2 1.
FT DOMAIN 618 712 SH2 2.
SQ SEQUENCE 724 AA; 81059 MW; 9D2BA8B6DB087098 CRC64;
Query Match 6.38; Score 146; DB 1; Length 724;
Best Local Similarity 22.38; Pred. NO. 0.026;
Matches 85; Conservative 44; Mismatches 148; Indels 104; Gaps 17;
QY 46 PLPDRRLAGVPGGECNSNNDYDEPE-----FOLLKAMPKMLLPARP----- 90

DB 95 PLPARRODPPPPGLTLPLPLPEQFSEPPDVAIPILVKLVFAIERTGLDSYRPPPAVRTDW 154
QY 91 -IQSESYADTRFQD-----MMAPLLPPKASSTSTORRDV/RMIDLEEVDPPTFFD 142
DB 155 SLSDVEQWMAALSDGVKGFLLALPAP-LVTEPAEAHRAHLEAAGPVGPALEPTLPL 213
QY 143 VRSQRFKFKYTKINKTP-LPPRPAP-LTLKKYQPL-----PPAPSESSAVFAPKP 193
DB 214 HNALTLR-FLDHLGVNAGRAPGPAVALGATGTEPILLRLRAPPPSPPGG-----APDG 268
QY 194 TPEEVORGRORSKANDSFVYLAEEESHQTRPSSCPSSNNTOKSPPAIASSSYMPK 253
DB 269 TEE-----TPPFPALVLEKLIQEHLEBQEVAPALPPKPKTKPA----- 308
QY 254 HSIQARDHGSQMOHCFAQRCQAASHSPMLPENTNSKPKPTKDEKDWQWNEYIGE 313
DB 309 -----PTGLANGSP-----PSLQDA---EWYKCD 330
QY 314 YSRQAVEDVLMKKNKDTFLVRDCSTKSKAEPVLVVFYGNKYVYKIRFLESNOQFALG 373
DB 331 ISREEVNKKL-RTPDGTFLVRDASKIQGE-YTLTRKGGN--NKLIVFHRDGHGFS 386
QY 374 TGLRGNMFDSYEDILIEHTY 394
DB 387 EPL-----TFCSVVDLTHYRH 403
RESULT 4
P85B_HUMAN
ID P85B_HUMAN STANDARD; PRT; 728 AA.
AC 000459;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
DE p85-beta subunit) (Ptdins-3-kinase p85-beta).
GN PIK3R2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98241181; PubMed=9582025;
RA Janssen J.W.G., Schleithoff L., Bertram C.R., Schulz A.S.;
RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase
RT p85beta subunit and HUMORF8, a putative deubiquitinating enzyme.";
RL Oncogene 16:1767-1772(1998).
CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X80907; CAA56868.1; -
CC HSSP: P23727; 2PNB.
DR Genew: HGNC:8980; PIK3R2.
DR MIM: 603157;
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000198; RhogAP.

DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PI3KINSEP85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00325; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 4 80 SH3.
 FT DOMAIN 125 258 RHO-GAP.
 FT DOMAIN 330 425 SH2 1.
 FT DOMAIN 622 716 SH2 2.
 SQ SEQUENCE 728 AA; 81624 MW; 80C2AF244977346B CRC64;

Query Match 6.3%; Score 146; DB 1; Length 728;
 Best Local Similarity 22.9%; Pred. No. 0.027;
 Matches 87; Conservative 48; Mismatches 147; Indels 98; Gaps 18;

OY 46 PLPDHRLNLAGVGGKCNNDYEDPEFQ--LLKAMPSPK-----ILPAPPIQE 93
 DB 95 PLPAPRPGAPRPGGLTLPDLPEQSPDPVAPPLVLKVLVEAIRTGLDSEHYRPLPAPR 154
 OY 94 SEYA-----DTRYFQDMMEAPLL-LP-----PKASVSTENQTDVMTQLEEVKDTF 140
 DB 155 TDMSLSDVDQMTALADGICSEFLALPAPLVTPEASAEARALAEAGVGPALPEPTL 214
 OY 141 KDVSQRKGFYKTKINPTLPPLPPR--PAI-TLPKTYOPL--PPAPPESSAYFAPKPT 194
 DB 215 PLHRLTLR-FLIQLGNVARRAPALGPVAVRNLGATFGLLRAPPPSSPPGAPPGS 273
 OY 195 FPEVGRGRQSAKDFSKVLGAEEESHQOTPESSCPSSNONTOKSPPAIASSVMPGKH 254
 DB 274 EP-----SPDFPALVLEKLQEHLEDEEVAPPALPPKPPAKPA----- 312
 OY 255 SIQARDHTGSMQHCPCQAQAASHSPMLPEYNTNSEKPPDTKDEKDWQNEWYIGY 314
 DB 313 -----PTVL-----ANGSP-----PSLQDA-----EYWGDI 335
 OY 315 SRQAVEDVLMKNDGTFLVDCSTKSKAEPYLVVFGNKVYVYKIRPLESNOQFALGT 374
 DB 336 SREEVNEKL-RDTPDGTFLVDAASKIGE-YTILTRKGN--NKLIVFHRDGHYGPSE 391
 OY 375 GLRGEMEDSVEDITEHTY 394
 DB 392 PL---TFCSVVDLINHYRH 407

RESULT 5
 P85B_RAT STANDARD; PRT; 722 AA.
 AC Q63788;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) (PtdIns-3-kinase p85-beta).
 GN PIK3R2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=96214979; Pubmed=8621382;
 RA Inukel K., Anal M., Vanbrede E., Hosaka T., Katagiri H., Funaki M.,

RA Fukushima Y., Ogihara T., Yazaki Y., Kiyuchi M., Oka Y., Asano T.;
 RT "A novel 55-KDa regulatory subunit for phosphatidylinositol 3-kinase
 RT structurally similar to p55PIK is generated by alternative splicing
 RT of the p85alpha gene.";
 RL J. Biol. Chem. 271:5317-5320(1996).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATE
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D64046; BAA10926.1; -.
 DR HSSP: P23727; 2PNB.
 DR InterPro: IPR001720; PI3KINase_P85.
 DR InterPro: IPR00198; RhogAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PI3KINSEP85.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00325; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 4 80 SH3.
 FT DOMAIN 125 257 RHO-GAP.
 FT DOMAIN 324 419 SH2 1.
 FT DOMAIN 616 710 SH2 2.
 SQ SEQUENCE 722 AA; 81328 MW; 1208368B9F6FC95 CRC64;

Query Match 6.2%; Score 143.5; DB 1; Length 722;
 Best Local Similarity 22.1%; Pred. No. 0.038;
 Matches 80; Conservative 39; Mismatches 120; Indels 123; Gaps 17;

OY 138 PTFKDVRSQR--FKGFKYTKINKTLPPLPPPAITLPKKYQDLPAP--PEES----- 185
 DB 58 PGFNEKTRQRGDFGTVEFLGPAVALAPGPR--PGCPPLPAPRDPGSESGHTLASL 114
 OY 186 SAYFAPKPTPEVGRGRQSAKDFSKVLGAEEESHQOTPESSCPSSNONTOK----- 239
 DB 115 AEQSPESPAPPI-----LVKLIEAIEQAEIDSEFYSRPELPAPRTDWSLSDLEQWDR 167
 OY 240 -----SPPAIASSSY-----MPGKHSIQAR---D 260
 DB 168 TTIYDAVKGFLPALPAVVPPEAASEAYRAMREYTGVGVLPEPTLPQLQALTLRLQL 227
 OY 261 HTGSMQH---CPQRCQAAS-HSPRNL-----PYENTNSEKPPDTK-----D 300
 DB 228 HLGNVARRAPSPATAVAHALASAFGLLRAPPGGEGDSEPAADPFVLLERLVQEHVD 287
 OY 301 EKD-----VMQNEWYIGYSGRAVEDVLMKENDGTG 332
 DB 288 EODTAPALPPKSKKVPAPATLANGSTSLQDAEYWGDISREEVNEKL-RDTPDGTG 346
 OY 333 LVNDCSTKSKAEPYLVVFGNKVYVYKIRPLESNOQFALGTGLRGEMEDSVEDITEHTY 392

```

Db      347 LVKDASSKIOGE-YTLTKRKGNN--NKLTKVHHDGCHGSESL-----TFCSVVELISHY 399
QY      393 TY 394
Db      400 RH 401

RESULT 6
FGR_FSVGR STANDARD: PRT: 545 AA.
AC P00544:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-kinase transforming protein FGR (EC 2.7.1.112).
GN V-FGR OR SRC-2.
OS Feline sarcoma virus (strain Gardner-Rasheed).
OC Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_Taxid=11775;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=84097512; PubMed=6318314;
RA Natarro G., Robbins K.C., Reddy E.P.;
RT "Gene product of v-fgr onc: hybrid protein containing a portion of
RT actin and a tyrosine-specific protein kinase.";
RL Science 223:63-66(1984).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FGR
CC POLYPROTEIN.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: X00255; CAA25063.1;
DR PIR: A00653; TVMYR.
DR HSSP: P00523; 2PTR.
DR InterPro: IPR004001; Actin.
DR InterPro: IPR004000; Actin_like.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00022; Actin; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00268; ACTIN; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00406; ACTIN; 1.
DR PROSITE: PS01132; ACTIN_ACT_LIKE; 1.
DR PROSITE: PS00001; SH2; 1.
KM Polyprotein; Tyrosine-protein kinase; Transferase; Phosphorylation;
KM ATP-binding; Oncogene; SH2 domain.
FT DOMAIN 23 157 ACTIN.
FT DOMAIN 167 264 SH2.
FT DOMAIN 286 539 PROTEIN KINASE.
FT NP_BIND 292 300 ATP (BY SIMILARITY).

```

```

FT BINDING 314 314 ATP (BY SIMILARITY).
FT ACT_SITE 405 405 BY SIMILARITY.
FT MOD_RES 435 435 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 545 AA; 61516 MW; 7F4742EB7A7C413E CRC64;

Query Match
Best Local Similarity 23.6%; Pred. No. 0.034; Length 545;
Matches 66; Conservative 30; Mismatches 96; Indels 88; Gaps 13;

QY 165 RPAITLPKKRYOPLPAPPESSAYF-----APKPEPEVGRGRORSA 207
Db 6 RPAYCPRPLPLPLPTAMEEVALYIDNGSMCKAGFADGAPRAVPEISYGRRHQCV 65
QY 208 -----KDSRYLGAEEESHQHTKPESSCPSSNONTQSPPIASSSYPGKHSTQAR-- 259
Db 66 MVMGQKD--SYVGDEAOSKR-----GILTLKY-PIEHGIVTWMD 102
QY 260 DHTGSMOHCPAORCOAAASHSPMLPYENTNSEK-P-DPTKPEK----- 302
Db 103 DMEKTHHTFYELDKRAPEHEVLL-----TEAPLNKANEKMTQIMETFNIPSNV 156
QY 303 ----DVMQNEWYIGEXSRQAVE-DVLMKKNKDGTEPLVRDCSTKSKAEPLYLV-----FY 352
Db 157 APVDSIOAEWEYFGKIGRKDAEROLLSPGNARGAFLVRESETTKGA--YSLSRDMDEAR 214
QY 353 GKKVYNVTKRFLPSNQDFALGGLGKNGEMFDSVEDIIEHY 392
Db 215 GDVKNHYKIRKLDTG-----GYITTRAQFNSVQELVQHY 249

RESULT 7
SCK_HUMAN STANDARD: PRT: 540 AA.
AC P98077; Q9NPL5; O60230;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein SCK (Fragment).
GN SCK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98175964; PubMed=9507002;
RA Nakamura T., Muraoka S., Sanokawa R., Mori N.;
RT "N-Shc and Sck, two neuronally expressed Shc adapter homologs. Their
RT differential regional expression in the brain and roles in
RT neurotrophin and Src signaling.";
RL J. Biol. Chem. 273:6960-6967(1998).
RN [2]
RP SEQUENCE OF 72-499 FROM N.A.
RA MEDLINE=95090462; PubMed=7527937;
RA Kavanaugh W.M., Williams L.T.;
RT "An alternative to SH2 domains for binding tyrosine-phosphorylated
RT proteins.";
RL Science 266:1862-1865(1994).
RN [3]
RP SEQUENCE OF 72-328 FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Brower A., Ramirez M.,
RA Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A.,
RA Ganes J., Dangnan L., Erler A., Christensen M., Georgescu A.,
RA Avila J., Liu S., Attix C., Andreise T., Trankiel M.,
RA Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R.,
RA Thomas P., Quan G., Krommiller B., Arellano A., Sanders C., Ow D.,
RA Nolan M., Trong S., Kobayashi A., Olsen A.S., Carraro A.V.;
RT "Sequence analysis of a 3.5 kb contig in human 19p13.3 containing a
RT serine protease gene cluster";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 149-540 FROM N.A.

```

RA Carlin L., Estivill X., Sunoy L., Escarceller M.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN LIVER, ALSO PRESENT IN BRAIN.
 CC -1- SIMILARITY: CONTAINS 1 PID DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: TO SHC TRANSFORMING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB001451; BAA25798.1; -
 DR EMBL: AC006124; AAC97988.1; -
 DR EMBL: AL360254; CAB96175.1; -
 DR HSSP: P29353; IMIL.
 DR MIM: 605217; -
 DR InterPro: IPR000050; PID_domain.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00640; PID; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRINTS: PR00629; SHCPIDOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00462; PTB; 1.
 DR SMART: SM00522; SH2; 1.
 DR PROSITE: PS01179; PID; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR SH2 domain.
 FT NON_TER 1 1
 FT DOMAIN 105 287 PID.
 FT DOMAIN 445 536 SH2.
 FT CONFLICT 72 72 A -> P (IN REF. 2).
 SQ SEQUENCE 540 AA; 57623 MW; 0D83700910B1A5DB CRC64;
 Query Match 6.0%; Score 140; DB 1; Length 540;
 Best Local Similarity 22.3%; Pred. No. 0.046;
 Matches 96; Conservative 55; Mismatches 185; Indels 94; Gaps 18;
 QY 13 GFGDLRFQWNLKNSWPLSLKAGRCRAVLRLPDHRRRLAGVPGGKCNNDY-- 69
 DB 169 GKSLRRAGNSISIHSTIDGLSVLRATQYI--ANHMHPSISASGD--TMDTDVAY 224
 QY 70 --EDPERQ---LKAMPKMLPAPRQIES--EYADTRFGQDMMEAR-LILPKKASYST 120
 DB 225 VAKDPINORACHILECEGL---AQSISTVGOAFELRFKQYLHSPKVALPPRLAGP 280
 QY 121 ERQTRDVRMQLQLE-----EVDKPTFKDVSQRKFGKRYKIKIKTPRPPRAPITL 170
 DB 281 EESWAGDEEDSLHNNYNSIPGKPPGLGLVDSRLALTPCALALDGGPSPSLRDGSL 340
 QY 171 PKKQPLPAPPESSAYFAKPTFEVQ---RGPRQSAKDFSRVLGAEESSHHTQKPE 227
 DB 341 PWDVGSTGTAPRGD-----YVQADARCP---PDHEHLVYNTQGLDAPRPE 384
 QY 228 SSCSSNONTQKSPPAISSYMGKHSIQARDHTGSMQHCPAORCAAMASHSPMLPYE 287
 DB 385 DS-----PKKDLFDWRPFEDALK---LHECSVAAGVTAAPLPLE 420
 QY 288 NTNSEKPD---PTKPDENVQNMWYIGEYSRAVEDYLKMKNDGTFVLVDCSTKSAE 344
 DB 421 DQWSPPTRRAPVAPTEQQLQOEWPYHGMRRAERKL---RADGGFLVDSVT--NPG 475
 QY 345 PYVLVVFNGNVVWKIRFLESNOQFALGTGLRCNEMDESVDIEHTYPIILLIDCKD 404
 DB 476 QYVLTLGHAGQPKHLLLVDP-----GVVTRKDVLFESISHLIDH-----LQNGQP 522
 QY 405 KAARRKOCYL 414
 DB 523 IVAASESLHL 532

RESULT 8
 ID 3BP2_HUMAN STANDARD; PRT; 561 AA.
 AC P78314; O15373; 000500; P78315;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SH3 domain-binding protein 2 (3BP-2).
 GN SH3BP2 OR 3BP2 OR RES4-23.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RC MEDLINE=98403881; PubMed=9734812;
 RX Hadano S., Ishida Y., Ikeda J.-E.;
 RA "The primary structure and genomic organization of five novel
 RT transcripts located close to the Huntington's disease gene on human
 RT chromosome 4p16.3".
 RL DNA Res. 5:177-186(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS), AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=98403881; PubMed=9734812;
 RA Bell S.M., Shaw M., Jou Y.-S., Myers R.M., Knowles M.A.;
 RT "Identification and characterization of the human homologue of
 RT SH3BP2, an SH3 binding domain protein within a common region of
 RT deletion at 4p16.3 involved in bladder cancer".
 RL Genomics 44:163-170(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RC TISSUE=tonsil;
 RX Gokemeijer J., Deligiannidis K.E., Ligris K., Ernst T.J.;
 RA "3BP2 binds to phosphatidylinositols: linking the hemopoietic tyrosine
 RT kinase c-FES to the cytoplasmic membrane in a phosphorylation
 RT dependent mechanism".
 RL Blood 88:473A-473A(1996).
 RN [4]
 RP VARIANTs CRM GLN-415; PRO-415; ARG-418; HIS-418; LEU-418; ARG-420 AND
 RP GLU-420.
 RX MEDLINE=21275962; PubMed=11381256;
 RA Ueki Y., Tizabi V., Santanna C., Fukai N., Maulik C., Garfinkle J.,
 RA Nimniya C., dokamari C., Peters H., Hbal M., Rhee-Morris L.,
 RA Doss J.B., Krelborg S., Olsen B.R., Reichemberger E.;
 RT "Mutations in the gene encoding c-Abl-binding protein SH3BP2 cause
 RT cherubism".
 RL Nat. Genet. 28:125-126(2001).
 CC -1- FUNCTION: BINDS DIFFERENTIALLY TO THE SH3 DOMAINS OF CERTAIN
 CC PROTEINS OF SIGNAL TRANSDUCTION PATHWAYS. BINDS TO
 CC PHOSPHATIDYLINOSITOLS; LINKING THE HEMOPOIETIC TYROSINE KINASE
 CC FES TO THE CYTOPLASMIC MEMBRANE IN A PHOSPHORYLATION DEPENDENT
 CC MECHANISM.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
 CC short form; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in a variety of tissues including
 CC lung, liver, skeletal muscle, kidney and pancreas.
 CC -1- DISEASE: Defects in SH3BP2 are the cause of cherubism (CRM), an
 CC autosomal dominant inherited syndrome. It is characterized by
 CC excessive bone degradation of the upper and lower jaws, which
 CC often begins around three years of age. It is followed by
 CC development of fibrous tissue masses, which causes a
 CC characteristic facial swelling.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U5386; AAB72034.1; -

DR EMBL: AB000462; BAA19119.1; -

DR EMBL: AB000463; BAA19120.1; -

DR EMBL: AF000936; AAB5973.1; -

DR Genew: HGNC:10825; SH3BP2.

DR MIM: 602104; -

DR MIM: 118400; -

DR InterPro: IPR001849; PH.

DR InterPro: IPR000980; SH2.

DR Pfam: PF00017; SH2; 1.

DR Pfam: PF00169; PH; 1.

DR ProDom: PD000093; SH2; 1.

DR SMART: SM00233; PH; 1.

DR SMART: SM00252; SH2; 1.

DR PROSITE: PS50001; SH2; 1.

DR PROSITE: PS50003; PH_DOMAIN; 1.

DR SH2 domain; SH3-binding; Alternative splicing; Disease mutation.

KW DOMAIN 26 130

FT SITE 201 210 SH3-BINDING.

FT DOMAIN 205 212 MISSING (IN SHORT ISOFORM).

FT DOMAIN 236 240 POLY-PRO.

FT DOMAIN 457 555 POLY-PRO.

FT SH2 81 97 SH2.

FT VARSPLIC 81 97 VMRAEETTSNNVEPFK -> QPRPOAALSQTEAGP (IN SHORT ISOFORM).

FT VARSPLIC 98 561 MISSING (IN SHORT ISOFORM).

FT VARIANT 415 415 R -> P (IN CRM).

FT VARIANT 415 415 /FTId=VAR_013257.

FT VARIANT 415 415 R -> Q (IN CRM).

FT VARIANT 418 418 /FTId=VAR_013258.

FT VARIANT 418 418 P -> H (IN CRM).

FT VARIANT 418 418 /FTId=VAR_013259.

FT VARIANT 418 418 P -> L (IN CRM).

FT VARIANT 418 418 /FTId=VAR_013260.

FT VARIANT 418 418 P -> R (IN CRM).

FT VARIANT 420 420 /FTId=VAR_013261.

FT VARIANT 420 420 G -> E (IN CRM).

FT VARIANT 420 420 /FTId=VAR_013262.

FT VARIANT 420 420 G -> R (IN CRM).

FT VARIANT 420 420 /FTId=VAR_013263.

FT CONFLICT 27 27 V -> L (IN REF. 3).

FT CONFLICT 224 224 H -> N (IN REF. 3).

FT CONFLICT 249 249 L -> R (IN REF. 3).

FT CONFLICT 251 251 A -> P (IN REF. 3).

FT CONFLICT 251 251 A -> R (IN REF. 3).

SO SEQUENCE 561 AA; 62243 MW; 69E6846A4F6D8F15 CRC64;

Query Match 5.98; Score 136.5; DB 1; Length 561;

Best Local Similarity 22.48; Pred. No. 0.06;

Matches 94; Conservative 44; Mismatches 138; Indels 143; Gaps 23;

QY 66 NNDEDEPEFOLKAMPKSKILPAPRIQSEYADT-----RYFQDMWAPILLPPKASVS 119

DB 171 NEDVHND-----EDDSLLEPDSPEPGLDALMPPVPP--FVP 210

QY 120 TEROTRDVMTQLEBVDPKFKVRSQRFKFKYTKINKTP-LPPRPATIL----- 170

DB 211 TPR-----KPAFSDMP-----RAHSFTSKGPGPLPPPPKGLPVDGLAAE 252

QY 171 -----PKKYQPLP--PAPPEE-SSAYFAPKPTFPFVQGGPROR----- 205

DB 253 DSKDPLPCPRAPPCPRVAPVPRRMSDPLSTMTAPGLRPPCFRESASPSPEPWPGH 312

QY 206 -SANDFSVLGAEEESHQTK-----PESSCPSSNON-----TOKSPATAS 246

DB 313 GACSTSSAIAIATATSRNCDLKSFLHSRGPPTSEPPVPAANKPKFLKIEEDPPRAA 372

QY 247 SSYPGKHSIARDHTGSMQHPAORCOAASHSPR---MLP-----YENTNS 291

DB 373 ---MGLFVPPVADRPALK-LPYPEAMARPAVLPRPEKPOLPHLQRCSPDGQSFSSSF 428

QY 292 EKP-DPFRPD-----EKDWQNEWYIGESRQAVEVLMK-----ENKDGTFIVR 335

DB 429 EKPRPSADADGCDSDSDYKVPPLPNSVFNNTSCVEHLFRATSPRGEPQDGLYICR 488

QY 336 DCSIKSKAEPYLVVE--YGNKYVNVKIRPLESNOQFALGTGLRCNENFSDVEDIIEHY 392

DB 489 NSSTKSGK---VLVWDETSNKRVRNYRI--FEKDSKF-----LGEVLVSVGSNVEHY 538

RESULT 9

BMX_HUMAN

ID BMX_HUMAN STANDARD; PRT; 675 AA.

AC P51813; Q12871; O60564;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cytoplasmic tyrosine-protein kinase BMX (EC 2.7.1.112) (Bone marrow kinase BMX) (Epithelial and endothelial tyrosine kinase) (ETK) (NTRK38).

DE BMX.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RP TISSUE=Bone marrow;

RP MEDLINE=95060827; PubMed=7970727;

RA Tamagnone L., Laitinen I., Mustonen T., Virtaneva K., Francis F., Muscatelli F., Allitalo R.P., Smith C.I., Larsson C., Allitalo K.;

RT "BMX, a novel nonreceptor tyrosine kinase gene of the BTK/ITK/TEC/TKX family located in chromosome Xp22.2.";

RL Oncogene 9:3683-3688(1994).

RN [2]

RP SEQUENCE FROM N.A., FUNCTION, AND INDUCTION.

RP TISSUE=Prostate;

RP MEDLINE=98188256; PubMed=9520419;

RA Qiu Y., Robinson D., Pretlow T., Kung H.-J.;

RT "Elk/Bmx, a tyrosine kinase with a pleckstrin-homology domain, is an effector of phosphatidylinositol 3'-kinase and is involved in interleukin 6-induced neuroendocrine differentiation of prostate cancer cells.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:3644-3649(1998).

RN [3]

RP SEQUENCE OF 536-599 FROM N.A.

RP TISSUE=Blood;

RA Fuortes M.;

RL Thesis (1994), Cornell University, U.S.A.

CC -1- FUNCTION: ACTIVITY IS REQUIRED FOR INTERLEUKIN 6 (IL-6) INDUCED DIFFERENTIATION. MAY PLAY A ROLE IN THE GROWTH AND DIFFERENTIATION OF HEMATOPOIETIC CELLS. MAY BE INVOLVED IN SIGNAL TRANSDUCTION IN ENDOCARDIAL AND ARTERIAL ENDOTHELIAL CELLS.

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein tyrosine phosphate.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN EPITHELIAL AND ENDOTHELIAL CELLS.

CC -1- INDUCTION: ACTIVATED BY IL-6 THROUGH PHOSPHATIDYLINOSITOL 3-KINASE (PI3-KINASE) PATHWAY. IT IS LIKELY THAT ACTIVATION OCCURS THROUGH BINDING OF PHOSPHOSITIDES TO THE PH DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).


```

DR EMBL: X83107; CAAS8169.1; -
DR EMBL: U08341; AAA17744.1; -
DR EMBL: AF045459; AAC08966.1; ALT_INIT.
DR HSSP: Q06187; 1B55.
DR Genew: HGNC:1079; BMX.
DR MIM: 300101; -
DR InterPro: IPR001562; BTK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000980; SH2.
DR Pfam: PF000169; SH2.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00779; BTK; 1.
DR PRINTS: PR00402; TECBTKDOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase; 1.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00107; BTK; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00232; SH2; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR Transferrase: Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; SH3 domain; SH2 domain.
FT DOMAIN 4 111 PH.
FT DOMAIN 196 289 SH3.
FT DOMAIN 296 392 SH2.
FT DOMAIN 417 675 PROTEIN KINASE.
FT N-BIND 423 431 ATP (BY SIMILARITY).
FT BINDING 445 445 ATP (BY SIMILARITY).
FT ACT_SITE 536 536 BY SIMILARITY.
FT MOD_RES 566 566 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DOMAIN 251 255 POLY-SER.
FT DOMAIN 282 285 POLY-SER.
FT DOMAIN 286 289 POLY-GLU.
FT CONFLICT 597 597 A -> S (IN REF. 3).
SQ SEQUENCE 675 AA; 78010 MW; CB2382A3BD02599 CRC64;

Query Match 5.7%; Score 133; DB 1; Length 675;
Best Local Similarity 24.08; Pred. No. 0.16; Indels 48; Gaps 11;
Matches 62; Conservative 41; Mismatches 107;

QY 166 PAILPKKQPLPPAPPESSAYFAPKPTPE-VQGRPR-----QSAKDFSVLGA--- 216
DB 141 PGCTLMAYANLHRAVNEKHRV---PTPPDRLKIPRAVPVLKMDAPSSSTTLAQYDN 196
QY 217 EEEHHOTKPESSCPSSNQNTQKSPRAIASS-----SYMGC-----KHSIQAR 259
DB 197 ESKKNYSSQPPSSSTSLAQYDSNKKIYGSQPNFMQYIPREDDPPDMWQYAKLSSSSSE 256
QY 260 DHTSMOHCAPQCAASHPRMLPYENTNSEKPDPTKPEKDVQWQNEWYIGESYQAV 319
DB 257 DVASSNKK-----ERNVNHSTKISWEPPSSSE-----EEVLDYDWFAGIISNOS 306
QY 320 EDVLAKENKDGTLVRCSTKSKAPVLYVF---YGNKYVNVKIRPLESNOQFALGTGL 376
DB 307 EQLLRQKGEKAGVVRN---SSQVGMVTVLSFSAVNDKGTGKVAHVHTNAEKKL--YL 361
QY 377 RGNEMFDSVEDIIEHTY 394
DB 362 AENYCFDSIPRLHYHQH 379

```

```

DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
DE p85-alpha subunit) (Pcdins-3-kinase p85-alpha) (PI3K).
GN PI3K1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBITaxID:10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS P85-ALPHA AND P55-ALPHA).
RC STRAIN:Wistar; TISSUE:Brain;
RX MEDLINE=96214979; PubMed=8621382;
RA Inukai K., Anal M., van Breda E., Hosaka T., Katagiri H., Funaki M.,
RA Fukushima Y., Ogihara T., Yazaki Y., Kikuchi M., Oka Y., Asano T.,
RT "A novel 55-kDa regulatory subunit for phosphatidylinositol 3-kinase
RT structurally similar to p55PIK is generated by alternative splicing
RT of the p85alpha gene."
RL J. Biol. Chem. 271:5317-5320(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).
RC TISSUE=Liver;
RX MEDLINE=97218222; PubMed=9065454;
RA Inukai K., Funaki M., Ogihara T., Katagiri H., Kanda A., Anal M.,
RA Fukushima Y., Hosaka T., Suzuki M., Shin B., Takata K., Yazaki Y.,
RA Kikuchi M., Oka Y., Asano T.;
RT "p85alpha gene generates three isoforms of regulatory subunit for
RT phosphatidylinositol 3-kinase (PI 3-kinase), p50alpha, p55alpha, and
RT p55alpha, with different PI 3-kinase activity elevating responses to
RT insulin."
RL J. Biol. Chem. 272:7873-7882(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).
RC TISSUE=Liver;
RX MEDLINE=97079666; PubMed=8921377;
RA Furman D.A., Cantley L.C., Carpenter C.L.;
RT "Structural organization and alternative splicing of the murine
RT phosphoinositide 3-kinase p85 alpha gene."
RL Genomics 37:113-121(1996).
CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: P85-ALPHA (SHOWN HERE), P55-
CC ALPHA AND P50-ALPHA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: THE P85-ALPHA ISOFORM IS WIDELY EXPRESSED.
CC EXPRESSION OF THE P55-ALPHA ISOFORM IS HIGHEST IN BRAIN AND
CC SKELETAL MUSCLE. THE P50-ALPHA ISOFORM IS ABUNDANT IN LIVER WITH
CC LOWER LEVELS IN BRAIN AND MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64045; BAA18932.1; -
DR EMBL: D64048; BAA18933.1; -
DR EMBL: U50412; AAC52846.1; -
DR EMBL: D78486; BAA24426.1; -
DR HSSP: P23727; 1BFI.
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000198; HOGAP.
DR InterPro: IPR000980; SH2.

```


DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RHO GAP; 1.
 DR PRINTS: PRO0678; PTKINASEP5.
 DR PRINTS: PRO0401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RHO GAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH3 domain; Repeat; Alternative splicing.
 FT DOMAIN 3 79 SH3.
 FT DOMAIN 333 428 SH2 1.
 FT DOMAIN 624 718 SH2 2.
 FT VARSPLIC 1 270 MISSING (IN ISOFORM P55-ALPHA).
 FT VARSPLIC 271 304 VTERPAAASDNTMLKAVELLISAEWSEKOPA -> MYT
 FT VARSPLIC 271 304 TVTWMEDLDIECAKTDINCCTDIMEFYEMDP (IN
 FT VARSPLIC 1 300 MISSING (IN ISOFORM P55-ALPHA).
 FT VARSPLIC 301 306 ROPAPA -> MHNLTQ (IN ISOFORM P50-ALPHA).
 SQ SEQUENCE 724 AA; 83531 MW; 95C65CF612873B84 CRC64.

Query Match 5.7%; Score 133; DB 1; Length 724;
 Best Local Similarity 22.2%; Pred. No. 0.18;
 Matches 88; Conservative 46; Mismatches 155; Indels 108; Gaps 19;

QY 38 GRCRAVLEPLPDHR--RNLAGVPGGKCNNDYEDPEFOLKAMP5MKTLPAPRIOESE 95
 DB 78 GRRK-ISPPTKPRPRPRPLPVAPGSSKTEA--DTEQPVLT-----PDIAEQ 121
 QY 96 YATRFQDMMEAPLLPRAVS-----TERQDVRMQLPEVDKPTF----- 140
 DB 122 FAPPD-----VAPPLIKLEALEKKGLECSIKYQSSNPALROLDCDDPSVDLV 176
 QY 141 --KDVRSOREKFGKYRKINKTPPPPPATLPKRYOPL-PPAPPESSAYFAVK---P 193
 DB 177 FDEHVLADAFKRY-----LADLPNPIYIPAVYNNEMMSLAQEVPSDDYIQLKLIRSP 230
 QY 194 TPEVVGGRPRQSAKDFRYLGADEESHQTKPESSC-----PSSNQNTQKSPPAI 244
 DB 231 NIHQYWLTLQYLLKHFFKLSQASSKNLLNARALSEIFSHVLFRRFAASDNT----- 284
 QY 245 ASSSYMPGKHSIOARDHTGSMQHPQRCQAASHSPRMPIYENTNEKRDPKP----- 299
 DB 285 -----HLKAVELLISAEMSEROPAPA-----LP-----KPPKPSIAN 319
 QY 300 ----DEKDVQNMENYIGEVSRQAVEDVLMKENKDGTEFLVRCSTKSKAEPIVLVVFY 355
 DB 320 NSMNNMNSLODAEMYMWDISREVENKLT-RDTADGTEFLVADASTKMGD-YTLTLKRG 377
 QY 356 VYVVKIRFLESNOQFALGTGLRGNEFDSVEDIIEHY 392
 DB 378 --NKLKIKIFHRDGKYGFSDEL-----TENSVELINHY 408

RESULT 11

ID GRAP_HUMAN STANDARD; PRT; 217 AA.
 AC 013588;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GRB2-related adaptor protein.
 DE GN GRAP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.

RX MEDLINE=96218119; PubMed=8647802;
 RA Feng G.-S., Ouyang Y.-B., Hu D.-P., Shi Z.-Q., Genz R., Ni J.;
 RT "Grap is a novel SH3-SH2-SH3 adaptor protein that couples tyrosine
 RT kinases to the Ras pathway.";
 RT J. Biol. Chem. 271:12129-12132(1996).
 CC -1- FUNCTION: COUPLE SIGNALS FROM RECEPTOR AND CYTOPLASMIC TYROSINE
 CC KINASES TO THE RAS SIGNALING PATHWAY.
 CC SUBUNIT: ASSOCIATES THROUGH ITS SH2 DOMAIN WITH LIGAND-ACTIVATED
 CC RECEPTORS FOR STEM CELL FACTOR (KIT) AND ERYTHROPOIETIN (EPOR).
 CC ALSO FORMS A STABLE COMPLEX WITH THE BCR-ABL ONCOPROTEIN. GRAP IS
 CC PRIMARILY THROUGH ITS N-TERMINAL SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE GRB2 / SEM-5 / DRK FAMILY.
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).

CC EMBL: U52518; AAC50541.1; -
 CC HSSP: 060631; IGBQ.
 CC Genew: HGNC:4562; GRAP.
 DR MTM: 604330; -
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 2.
 DR PRINTS: PRO0452; SH3DOMAIN.
 DR ProDom: PD000066; SH3; 2.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 2.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 2.
 KW SH2 domain; SH3 domain; Repeat.
 FT DOMAIN 1 58 SH3 1.
 FT DOMAIN 60 152 SH2.
 FT DOMAIN 156 215 SH3 2.
 SQ SEQUENCE 217 AA; 25336 MW; 09FEC2F3BAC0FA8 CRC64;

Query Match 5.7%; Score 132.5; DB 1; Length 217;
 Best Local Similarity 27.7%; Pred. No. 0.048;
 Matches 36; Conservative 24; Mismatches 41; Indels 29; Gaps 4;

QY 293 KPDPKPDQDMQNMENYIGEVSRQAVEDVLMKENKDGTEFLVRCSTKSKAEPIVLVVFY 352
 DB 56 KPRP-----WYSGRISQLAEELIKRNHGLAIRL-SESSGPESSVYNY 101
 QY 353 GNRVYVYKIRFLESNOQFALGTGLRGNEFDSVEDIIEHYTYPIILLIDGKRAARRKQC 412
 DB 102 GDQVHFVKYLRASGKVEL-----WEKFNLSLVELVDYFRTTYI-----AKKRQI 146
 QY 413 YLTQPLPLAR 422
 DB 147 FLRDEEPLK 156

RESULT 12

ID CRK_DROME STANDARD; PRT; 271 AA.
 AC 09XVMO;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adapter molecule ctk.
 DE GN CRK OR CG1387.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

CC Insecta:terrygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila;
OX NCBI_taxid=7227;
RN [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RP
RP TISSUE=EMBRYO;
RC MEDLINE=99173888; PubMed=10072777;
RX Galletta B.J., Niu X.-P., Erickson M.R., Abmayr S.M.;
RT "Identification of a Drosophila homologue to vertebrate Crk by
RT interaction with Mbc.";
RL Gene 228:243-252(1999).
[2]
RN
RN SEQUENCE FROM N.A.
RP
RP STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blasei R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkeva D., Botchan M.R., Bouck Y., Brockstein P., Brotlier P.,
RA Butliss K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.C., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman J.U., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA Jatalin M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laaslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclel J.M.,
RA Palazzolo M., Plutman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svyrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC
CC -1- FUNCTION: ADAPTER PROTEIN WHICH INTERACTS WITH C-TERMINAL PORTION
CC OF Mbc, HOMOLOG OF HUMAN DOCK180. MAY PLAY A ROLE IN CELLULAR
CC PROCESSES THROUGHOUT DEVELOPMENT.
CC
CC -1- TISSUE SPECIFICITY: EMBRYONIC ZYGOTIC EXPRESSION IS SEEN IN
CC IMAGINATING PRESUMPTIVE MESODERM AND ECTODERMALLY DERIVED TISSUES
CC DURING GASTRULATION. AT STAGE 8, EXPRESSION IS ALSO SEEN IN
CC ANTERIOR AND POSTERIOR MIDGUT AND CEPHALIC FURROW. BY STAGE 9,
CC EXPRESSION IS HIGHEST IN VISCERAL MESODERM OF ANTERIOR AND
CC POSTERIOR MIDGUT, VENTRAL NERVE CORD AND SOMATIC MESODERM.
CC
CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY
CC THROUGHOUT EMBRYOGENESIS, DECLINES DURING LARVAL STAGES AND
CC REAPPEARS DURING PUPATION.
CC
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC
CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
CC
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC  entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE003844; AAF59362.1; -
DR  EMBL; AF112976; AAD28428.1; -
DR  HSSP; 064010; 1CKA.
DR  FLYBase; FBgn0024811; Crk.
DR  InterPro; IPR000980; SH2.
DR  InterPro; IPR001452; SH3.
DR  Pfam; PF00017; SH2; 1.
DR  Pfam; PF00018; SH3; 1.
DR  PRINTS; PRO0401; SH2DOMAIN.
DR  PRINTS; PRO0452; SH3DOMAIN.
DR  PRODom; PD000066; SH3; 1.
DR  PRODom; PD000093; SH2; 1.
DR  SMART; SM00252; SH2; 1.
DR  SMART; SM00326; SH3; 2.
DR  PROSITE; PS50001; SH2; 1.
DR  PROSITE; PS50002; SH3; 1.
KW  SH2 domain; SH3 domain; Repeat; Developmental protein.
FT  DOMAIN 12 114 SH2
FT  DOMAIN 117 165 SH3 1.
FT  DOMAIN 220 259 SH3 2.
SQ  SEQUENCE 271 AA; 31205 MW; D1BAFE43150932DC CRC64;

Query Match 5.7%; Score 132.5; DB 1; Length 271;
Best Local Similarity 35.6%; Pred. 0.062;
Matches 37; Conservative 13; Mismatches 49; Indels 5; Gaps 2;

OY 306 ONEWIGEXSQAVEDVLMKENKOGTFVLRCCSTKSKAEPLVLYVFYGNKYNNKIRPLE 365
DB 9 RNSWTFGMSRQADATEVLMNERKGVFLVRD--SNSIAGDYLVCREDTKVSNYIINKVQ 66
OY 366 SNOQFALCTGLRGNEFDSVEDIIEHYTPFILLIDGKDAARR 409
DB 67 QODQIVYRI---GDQSFNDLPKLLFTVYLHYLDTPPLKPRACRR 107

RESULT 13
ABP1_YEAST STANDARD; PRT; 592 AA.
ID ABP1_YEAST
AC P15891.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Actin binding protein.
GN ABP1 OR YCR088W OR YCR88W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90136906; PubMed=2405279;
RA Dublin D.G., Mulholland J., Zhu Z., Botstein D.;
RT "Homology of a yeast actin-binding protein to signal transduction
RT proteins and myosin-I.";
RL Nature 343:288-290(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Frontali L., Grisanti P.;
RL submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN THE SPATIAL ORGANIZATION OF CELL
CC SURFACE GROWTH. AN OVERPRODUCTION OF ABP1 CAUSES THE ASSEMBLY OF
CC THE CORTICAL ACTIN SKELETON AT INAPPROPRIATE SITES ON THE CELL
CC SURFACE, RESULTING IN DELocalIZED SURFACE GROWTH.
CC -1- SUBCELLULAR LOCATION: CORTICAL CYTOSKELETON.
CC -1- SIMILARITY: TO SIGNAL TRANSDUCTION PROTEINS AND MYOSIN I.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X51780; CAA36075.1; -
DR EMBL: X59720; CAA42253.1; -
DR PIR: S19503; LBBY.
DR HSSP: P29355; ISEM.
DR SGD: S0000684; ABP1.
DR InterPro: IPR002108; Actbind_cofin.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00241; cofilin_ADF; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00102; ADF; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
DR Cytoskeleton; Actin-binding; SH3 domain; Repeat.
FT SIMILAR 1 142 TO YEAST COFILIN.
FT DOMAIN 86 97 ACTIN-BINDING (POTENTIAL).
FT DOMAIN 532 592 SH3.
FT DOMAIN 200 575 3 x 10 AA APPROXIMATE REPEATS.
FT REPEAT 200 209 1.
FT REPEAT 436 445 2.
FT REPEAT 566 575 3.
FT REPEAT 575 58 L -> S (IN REF. 1).
FT CONFLICT 312 312 K -> I (IN REF. 1).
FT CONFLICT 592 AA; 65576 MW; 395235107049494A CRC64;
SQ SEQUENCE

Query Match 5.7%; Score 132.5; DB 1; Length 592;
Best Local Similarity 21.4%; Pred. No. 0.15;
Matches 54; Conservative 44; Mismatches 99; Indels 55; Gaps 11;

QY 117 SVSTERQTRDVRMTQLEVDKPTFKVRSQREKGFYTKINKTPL-----PPPR---PAT 168
DB 329 NAKTKAEAKRPEVDEPEGEPEVDKDKLS-KFGLAASEKEEEMENKFAFPKKSSPTL 387
QY 169 TLPRKY-QPLPPAPRESSAYFAKPTFPEVQGPQRQSKDNRVLGAEEESHQTKPE 227
DB 388 ISKRPKSKQEPKAEAE--QPKTDYKKIG-----NPLPGMHIEDNEEPE 433
QY 228 SS-----CPSSN-----ONTQKSPPAISSYMPGKHSIOARDHTGSM 265
DB 434 ENDDDDDEDAAPPLPSRNVAASGAPVQKEPEGEELAPS--LFSRNITPAKQEEAP 491
QY 266 QHCPCACQAAASHSPRLPYENTNSEKDPPTK-PDEKDVQWQWYIGEYSRQAVEDYLM 324
DB 492 EQAPEEIEEAEAEAPQLPSRSSAAPPPPRRATPEKKRKNPMTATAEYDYAAD--- 548
QY 325 KENKQSTFLVRD 336
DB 549 ---NELTFEEND 557

RESULT 14

PSC_DROME STANDARD; PRT; 1603 AA.

AC P35820;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Posterior sex combs protein.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP
RX SEQUENCE FROM N.A. PubMed=1833647;

RA Brunk B.P., Martin E.C., Sharp E., Adler P.N.;
RT "Drosophila genes Posterior Sex Combs and Suppressor two of zeste
RT encode proteins with homology to the murine bmi-1 oncogene";
RL Nature 353:351-353(1991).
CC -!- FUNCTION: THE POLYCOMB GROUP (PC-G) GENES ARE NEEDED TO MAINTAIN
CC EXPRESSION PATTERNS OF THE HOMEOBOX SELECTOR GENES OF THE
CC ANTENNAPEDIA (ANTP-C) AND BITHORAX (BX-C) COMPLEXES, AND HENCE FOR
CC THE MAINTENANCE OF SEGMENTAL DETERMINATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X59275; CAA41965.1; -
DR PIR: S17983; S17983.
DR FlyBase: FBgn0005624; Psc.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00518; ZF_RING_1; 1.
DR PROSITE: PS50089; ZF_RING_2; 1.
KW Zinc-finger; Developmental protein; DNA-binding; Nuclear protein.
FT DOMAIN 47 53 POLY-THR.
FT DOMAIN 83 88 POLY-THR.
FT DOMAIN 91 98 POLY-THR.
FT DOMAIN 145 152 POLY-THR.
FT DOMAIN 184 202 POLY-SER.
FT ZN_RING 265 304 RING-TYPE.
FT DOMAIN 642 651 POLY-SER.
FT DOMAIN 1066 1069 POLY-GLY.
FT DOMAIN 1185 1189 POLY-PRO.
FT DOMAIN 1214 1217 POLY-PRO.
FT DOMAIN 1391 1396 POLY-PRO.
FT DOMAIN 1458 1461 POLY-ALA.
FT DOMAIN 1517 1520 POLY-GLY.
SQ SEQUENCE 1603 AA; 169999 MW; 77024f409736473 CRC64;

Query Match 5.7%; Score 132; DB 1; Length 1603;
Best Local Similarity 21.2%; Pred. No. 0.52;
Matches 80; Conservative 51; Mismatches 148; Indels 98; Gaps 18;

QY 26 KNSWPSLSSAKGRCR-----AVLEPLDRHRLAGVPG-----GEKCSNNDYEDPEFOL 76
DB 641 KSSSSSSSSSGERKRKSPPLVPLTIRTERIMSPSGVSTLSPRVTSQAFSDPKSEF 700
QY 77 LKAMP-----SKILPAR-IQSEVADR--YFQD--WMEAPLLPPK 115
DB 701 LKSFALKPIKVKVSPERTLNRAITPPSPVSQASPKSGNNLDDSLIMKPPSCMPK 760
QY 116 ASYSTERQTRD---VMTQLEVDKPT--FK---DVSORFQKGFYTKINKTPLPP-R 165
DB 761 SIASSKRSKKEPVKAVSKKQKLSPLPTVPDKIRLPYTNNGSSSTASPKIEKPLMPPAK 820
QY 166 PAITLPKKYQPL-----PPAPRESSA--YEAP-----KPTPEVQR-GPRQSA 207
DB 821 PPMILAPRKIQPSAQFAPRPPPIHHAGVMSAPGNRTPIAKRYOPLPKASRPFPANIP 880
QY 208 KDFSRYL-----GAEEESHQTKPRESSCSSNQNQKSPPAISSYMGKH 254
DB 881 NDVNRLLDKAGTEIKSIGGSVENSNSAOKPHLYGPKG--ETKMGPPALPATTPSGGNK 938
QY 255 SIQARDITGSMOHC-----AQRCAAAASHSPMLPYENTNSEKDPPTK 298
DB 939 NVGKQAGNLPMSAPPNKNGSSNNYLNALFNSMKNCKQKAPPGCGRTPMYTPNSITYSPSS 998
QY 299 PDEKDVQWQWYIGEYS 315

DB 999 PQ-----YVPSYN 1006

RESULT 15

DB P85B_MOUSE STANDARD: PRT: 722 AA.

AC 008908;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase p85-beta subunit) (Ptdins-3-kinase p85-beta).

GN PI3K2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.

OX NCBI_Taxid=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=NIH Swiss;

RX MEDLINE=98241181; PubMed=9582025;

RA Janssen J.W.G., Schleithoff L., Bartram C.R., Schulz A.S.;

RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subunit and HUNKOR8, a putative deubiquitinating enzyme.";

RL Oncogene 16:1767-1772(1998).

CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE KINASERS, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.

CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY) SUBUNIT.

CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.

CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Y13569; CAI73903.1; ..

DR HSSP: P23727; ZPMB.

DR MGD: MGI:1098772; PI3K2.

DR InterPro: IPR001720; PI3kinase_P85.

DR InterPro: IPR000198; RHO-GAP.

DR InterPro: IPR000980; SH2.

DR InterPro: IPR001452; SH3.

DR Pfam: PF00017; SH2; 2.

DR Pfam: PF00018; SH3; 1.

DR Pfam: PF00620; RHO-GAP; 1.

DR PRINTS: PRO0678; PI3KINASEP85.

DR PRINTS: PRO0401; SH2DOMAIN.

DR PRODOM: PD000066; SH3; 1.

DR PRODOM: PD000093; SH2; 2.

DR SMART: SM00324; RHO-GAP; 1.

DR SMART: SM00252; SH2; 2.

DR SMART: SM00326; SH3; 1.

DR PROSITE: PS50001; SH2; 2.

DR PROSITE: PS50002; SH3; 1.

DR SH3 domain; SH2 domain; Repeat.

FT DOMAIN 4 80 SH3.

FT DOMAIN 125 257 RHO-GAP.

FT DOMAIN 324 419 SH2 1.

FT DOMAIN 616 710 SH2 2.

SO SEQUENCE 722 AA; AB4D49AF30CEC567 CRC64;

Query Match 5.68; Score 130.5; DB 1; Length 722;

Best Local Similarity 21.58; Pred. No. 0.26; Mismatches 149; Indels 129; Gaps 20;

Matches 88; Conservative 44;

OY 53 NLACYPG-GEKNSNNDYEDP--EF-----QLKAMP-----SMKILPAPRIQSEYADTRF 102

DB 53 NVGMPGFENERRGDEFGTYVEFLGPVALAPGPBPRGPPLPAPRDGSSSEG-HIL 111

OY 103 QDME---APLLLP-----KASVTERQTR-----DVRMTOLEEVKPKPFK 141

DB 112 PLAEQFSPDPAPPIVLKVLVAIEQAELDSQCYKPELPATRTDMSLDLEQMDRTALY 171

OY 142 DVRSQRFKGF-----KYTKINKT-----PLPPPAITLPPKKYQ 175

DB 172 DA-----VKGFLALPAAVVTPEAAAEAVFALREVAGPVGLVEPPTLLPLHQALTRELLQ 227

OY 176 PL---PPAPPESSAVFAKPTF-----PEVGRGRQRAKQESVLAGEEESHQT 224

DB 228 HIGVARRAPSPDTIVHALASAFGLLRIPSGSGEGSEVPDFPVLLERLVQEHVE 287

OY 225 KPSSCPSSNONTOKSPPAIASSSYMPGKHSIOARDHTGSMQHCPCARCOAAASHSPML 284

DB 288 EQDAAPALPPKPSKAKPA----- 306

OY 285 PYENTNSEKPDPTKPDKDWONENYIGEYSKQAVEDYLMKENKGTFLVRDCSTKSKAE 344

DB 307 PTALANGSP---PSLQDA---EWYWGDISREEVNERL-RDTPDGTFLVRDASSKIOGE 358

OY 345 PYLVVVFYGNKYVYKIRPLESNOQFALGTGRGNEMFDSVDIIEHTY 394

DB 359 -YTLIRKGN--NKLIVFHRDGHGFSERL-----TCSYVELLSYRHH 401

Search completed: April 21, 2003, 12:36:51

Job time : 13.6547 secs

GenCore version 5.1.4.D5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:14:41 ; Search time 50.706 Seconds

(without alignments)
5933.817 Million cell updates/sec

Title: US-09-856-061-3

Perfect score: 2020
Sequence: 1 ttccagacttcagctgcgc.....catgcacacacagaatt 1129

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-Q=/gen2.1/USPTO-spool/US09856061/runat_21042003_113018_3728/app_query.fasta.1.3150
-DB=A.Geneseq_101002 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061.ecgn.1.1.209 @runat_21042003_113018_3728 -NCP=6 -ICPU=3
-NO_XLPRX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMECUT=120
-WARN_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2020	100.0	376	22	AA81071	Human mast cell-st
2	2009	99.5	428	23	AAE22609	Human MIST splice
3	2009	99.5	443	23	AAE22608	Human MIST protein
4	1694	83.9	353	23	AAE22610	Human MIST splice
5	1132	56.0	435	22	AA81070	Murine mast cell-s
6	349	17.3	84	21	AA603152	Human secreted pro
7	290.5	14.4	449	19	AAW59866	Amino acid sequenc
8	290.5	14.4	456	22	AA893406	Human protein sequ
9	287.5	14.2	456	19	AAW59865	Amino acid sequenc
10	273	13.5	457	19	AAW59867	Amino acid sequenc
11	238	11.8	46	22	AB838147	Peptide #5653 enco
12	238	11.8	46	22	AB823554	Protein #5353 enco
13	238	11.8	46	22	AAW58783	Human brain expres
14	238	11.8	46	22	AAW71296	Human bone marrow
15	238	11.8	46	22	AAW18973	Peptide #5407 enco
16	238	11.8	46	22	AAW31575	Peptide #5612 enco
17	238	11.8	46	23	ABG41096	Human peptide enco
18	236	11.7	95	20	AAW70587	Human SLP-76, SH2 d
19	214	10.6	40	22	AB839614	Peptide #7120 enco
20	214	10.6	40	22	AB839615	Peptide #7121 enco
21	214	10.6	40	22	AAW60320	Human brain expres
22	214	10.6	40	22	AAW60321	Human brain expres
23	214	10.6	40	22	AAW72853	Human bone marrow
24	214	10.6	40	22	AAW72854	Human bone marrow
25	214	10.6	40	22	AAW33179	Peptide #7216 enco
26	214	10.6	40	22	AAW33180	Peptide #7217 enco
27	214	10.6	40	23	ABG42791	Human peptide enco
28	214	10.6	40	23	ABG42792	Human peptide enco
29	170	8.4	503	22	ABG22854	Novel human diagno
30	167	8.3	945	22	AB866077	Drosophila melanog
31	166	8.2	30	22	AB838148	Peptide #5654 enco
32	166	8.2	30	22	AB823555	Protein #5354 enco
33	166	8.2	30	22	AAW58784	Human brain expres
34	166	8.2	30	22	AAW71297	Human bone marrow
35	166	8.2	30	22	AAW18974	Peptide #5408 enco
36	166	8.2	30	22	AAW31576	Peptide #5613 enco
37	166	8.2	30	23	ABG41097	Human peptide enco
38	155	7.7	724	16	AAW85784	Human GRB-1, Homo
39	151.5	7.5	759	22	ABG21983	Novel human diagno
40	150.5	7.5	724	13	AAW26060	Growth Factor Rece
41	150.5	7.5	724	21	AAW97143	Human PI3K p85 pol
42	150.5	7.5	724	22	AAW67442	Amino acid sequenc
43	150.5	7.5	724	22	AAW67621	Amino acid sequenc
44	150.5	7.5	724	23	ABG60298	Human lymphoma ass
45	150.5	7.5	968	19	AAW85011	p85alpha-green flo

ALIGNMENTS

RESULT 1
ID AAB81071 standard; Protein: 376 AA.

AC AAB81071:
XX 25-JUN-2001 (first entry)

DE Human mast cell-specific immunoreceptor signal transducer.

KW Mast cell: signal transduction; human; allergic disease; MIST;

XX mast cell-specific immunoreceptor signal transducer.

OS Homo sapiens.

PN JP3146204-B1.

XX 12-MAR-2001.
XX

XX Claim 11, Fig 5, 171pp: English.

PS The present invention relates to novel mast cell immunoreceptor signal

CC transducer (MIST) proteins and polynucleotides encoding such proteins.

CC MIST sequences of the invention are useful for preventing, treating or

CC ameliorating a medical condition in mammalian subject. They are useful

CC for treating an immune disorder involving hyperactivity of B- or T-

CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-

CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful

CC as targets for therapeutic intervention in immune cell disorders and

CC inflammatory indications, for diagnosis and/or screening of disorders

CC or diseases associated with expression of MIST, for screening for

CC antagonists or inhibitors of the interaction of MIST with cellular

CC signalling components. They are used in assays that detect activation

CC or induction of various B and T-cell-related neoplasms or cancers.

CC Sequences of the invention are also used in gene therapy. The present

CC sequence is human MIST splice variant protein from clone #7.

XX

Sequence 428 AA:

Alignment Scores:

Pred. No.:	5,42e-195	Length:	428
Score:	2009.00	Matches:	374
Percent Similarity:	99.73%	Conservative:	1
Best Local Similarity:	99.47%	Mismatches:	1
Query Match:	99.46%	Indels:	0
DB:	23	Gaps:	0

US-09-856-061-3 (1-1129) x AAE22609 (1-428)

OY 1 TTCGAGAACTTCAGTCTGCCAAAAAAGAGTCATGCGCTCGCATATAGTCCACAGGC 60

DB 19 PheGlnAsnPheserLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGly 38

OY 61 CAGTACCGAGAGATGAACAGCCCTCTCTAGACTGGGAAAGAACTTCTGCACTCTG 120

DB 39 GlnTyrGlnArgMetAsnLysProLeuLeuAspTrpGlnArgAsnPhesAlaValLeu 58

OY 121 GATGAGCAAAAGGCCACAGTATGATGATGATGATGATGATGATGATGATGATGATG 180

DB 59 AspGlyAlaLysGlyHisSerAspAspTrpAspAspProGlnLeuArgMetGlnGlu 78

OY 161 ACATGGCAGTGAATTAATTTTACCAGCCGCGCTATTAAGAACTTGAATATGACAT 240

DB 79 ThrTrpGlnSerIleLysIleLeuProAlaArgProIleLysGlnSerLysValAsp 98

OY 241 ACACACTATTTCAAGCTTGCAATGACACTCCCTTCGCTTACAGACACAGACCTCTATC 300

DB 99 ThrHisTyrPheLysValAlaMetAspThrProLeuProLeuAspThrArgThrSerIle 118

OY 301 TCCATTGGACAGCCGACCTGGAACACACAGAGAGGTTGGAAGAGTGGAACAACTT 360

DB 119 SerIleGlyGlnProThrTrpAsnThrGlnThrArgLeuGlnArgValAspLysProIle 138

OY 361 TCCAGGAGCGTCAGAGCCAAAACATTAAGAGAGATGCAATCCGTAAAGAAAGCAAGATT 420

DB 139 SerLysAspValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLysIle 158

OY 421 CCTTATACCACTCTCGGCTCTCATACACTTCCGAGAGAACTACCACTTGGCCCT 480

DB 159 ProLeuProProProArgProLeuIleThrLeuProLysLysTyrGlnProLeuProPro 178

OY 481 GAGCGGAGAGAGAGAGGACCTTATCTCAGAGACACACTTCCCAAGTCCAGGGA 540

DB 179 GlnProGlnSerSerArgProProLeuSerGlnArgHisThrPheProGlnValGlnArg 198

OY 541 ATGCCAGTCAGATTAAGGACTTAAGGACTTAAGGAGTCTTGAAGCAGAAAAAGTTCT 600

DB 199 MetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnAlaIleLysValPro 218

OY 601 CATACACAGAGAGAGCGTGAATCAACTCATCTGTGTTAGAAAACCAAAATATCTCAAGAGATT 660

DB 219 HisAsnGlnArgLysProGlnSerThrHisLeuGlnAsnGlnAsnThrGlnGluIle 238

OY 661 CCACCTTGCCATTAGACAGTCTTCATTCACAGACAGAACACAGTGTGCAAAACAGAGAT 720

DB 239 ProLeuAlaIleLysSerSerSerSerPheThrThrSerAsnHisSerValGlnAsnArgAsp 258

OY 721 CATAGAGGAGGACATCAGCCCTGTTCTCTCAGAGATGCGACCTCCACAGCTGCAGC 780

DB 259 HisArgGlyGlyMetGlnProCysSerProGlnArgCysGlnProProAlaSerCysSer 278

OY 781 CCTCAGCAAAATATATCTGCGCTATTAATACACAGCTGAGACACACTTCCCAAAAG 840

DB 279 ProHisGlnAsnIleLeuProTyrLysTyrThrSerTrpArgProPheProLysArg 298

OY 841 TCTGATAGAAAGATGCTCAGACACATGAATGATGATGAGAAATACAGCCGAGCA 900

DB 299 SerAspArgLysAspValGlnHisAsnGlnTyrTrpIleGlyGlnTyrSerArgGlnAla 318

OY 901 GTGGAAGAGGCAATTCATGAGAGAGAAAGATGCTAGTTCTTGTCGAGATTGTTC 960

DB 319 ValGlnGlnAlaPheMetLysGlnLysAspGlySerPheLeuValArgAspCysSer 338

OY 961 ACAAATCCAGAGAGAGCCCTATGTTTGGCTGTGTTTATGAGACAAAGCTTACAT 1020

DB 339 ThrLysSerLysGlnGlnProTyrValLeuAlaValPheTyrGlnAsnLysValTyrAsn 358

OY 1021 GTAAAAATCCGCTTCTCGAGAGAAATCAGAGTTTGCCTCGGAGCAGACTCGAGAGA 1080

DB 359 ValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlyThrGlyLeuArgGly 378

OY 1081 GATGAGAACTTGAATTCATGATGAGAGACATCATGCACTACAGAAAT 1128

DB 379 AspGlnLysPheAspSerValGlnAspIleIleGlnHisTyrLysAsn 394

RESULT 3

ID AAE22608 standard; Protein: 443 AA.

XX AAE22608;

AC 26-JUL-2002 (first entry)

XX

XX Human MIST protein #1.

DE

XX Human: mast cell immunoreceptor signal transducer; MIST: immune disorder;

KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;

KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.

OS

XX Homo sapiens.

XX

XX Key Location/Qualifiers

FT Modified-site 84

FT Modified-site 111

FT /label= Tyrosine-phosphorylation_binding_site

FT Domain 306..311

FT /label= SH3-binding-proline-rich_motif

FT Domain 324..407

FT /label= SH2_domain

PN WO200226986-A2.

PD 04-APR-2002.

XX

PF 28-SEP-2001; 2001WO-US30593.

XX

PR 29-SEP-2000; 2000US-237030P.

XX

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX

PI Perez-Villal JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;

XX WPI, 2002-372126/40.

DR N-PSDB: AAD35800.
 XX
 PT New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
 PS
 PS Claim 11: Fig 3: 171pp: English.
 XX
 CC The present invention relates to novel mast cell immunoreceptor signal
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.
 CC Sequences of the invention are useful for preventing, treating or
 CC ameliorating a medical condition in mammalian subject. They are useful
 CC for treating an immune disorder involving hyperactivity of B- or T-
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
 CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
 CC as targets for therapeutic intervention in immune cell disorders and
 CC inflammatory indications, for diagnosis and/or screening of disorders
 CC or diseases associated with expression of MIST, for screening for
 CC antagonists or inhibitors of the interaction of MIST with cellular
 CC signalling components. They are used in assays that detect activation
 CC or induction of various B and T-cell-related neoplasms or cancers.
 CC Sequences of the invention are also used in gene therapy. The present
 CC sequence is human MIST protein.
 CC
 XX Sequence 443 AA:
 XX
 Alignment Scores:
 Pred. No.: 5 5e-195 Length: 443
 Score: 2009.00 Matches: 374
 Percent Similarity: 99.73% Conservative: 1
 Best Local Similarity: 99.47% Mismatches: 1
 Query Match: 99.46% Indels: 0
 DB: 23 Gaps: 0
 US-09-856-061-3 (1-1129) x AAE22608 (1-443)
 QY 1 TTCGAGAGCTTCAGTTCGCAAAAACAGAGTCATGCGCTGCATCAATAGTCCACAGGC 60
 DB 34 PheGlnAsnPheserLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGly 53
 QY 61 CAGTACGAGAGGATGAGCAAGCCCTCTAGACTGGGAAAGAACTTGGTCGACGTCTG 120
 DB 54 GlnTyrGlnArgMetAsnLysProLeuLeuAspTrpGlnArgAsnPheAlaAlaValLeu 73
 QY 121 GATGAGCAAAAGCCACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 DB 74 AspGlyAlaLysGlyHisSerAspAspAspAspAspAspAspAspAspAspAspAsp 93
 QY 181 ACATGAGAGTCGATTAATTTTACAGCCCGGCTTAAGGAATCTGAATATGACAGAT 240
 DB 94 ThrTrpGlnSerIleLysIleLeuProAlaArgProIleLysGlnSerGlnTyrAlaAsp 113
 QY 241 ACACACATTTTCAAGTTGCAATGACATCCCTCCGTTAGACACAGAGCTCTATC 300
 DB 114 ThrHisTyrPheLysValAlaMetAspTrpTrpLeuProLeuAspTrpArgThrSerIle 133
 QY 301 TCCATTTGAGAGCCGAGCTGGAACACAGAGAGTTGGAAGAGTGGACAAACCATT 360
 DB 134 SerIleGlyGlnProTrpTrpAsnThrGlnThrArgLeuGlnValAspLysProIle 153
 QY 361 TTCAGGAGCTCAGAACCCAAACATTAAGAGAGATCCCTTAAGAAAGAACAGATT 420
 DB 154 SerLysAspValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLysIle 173
 QY 421 CCTTACACCTCTCGGCTCTATACACTTCCGAAGAGTACCAACCTTCCCTCCCT 480
 DB 174 ProLeuProProProArgProLeuIleThrLeuProLysTyrGlnProLeuProPro 193
 QY 481 GAGCGGAGAGCAGCAGCCACCTTATCTCAGAGACACACCTTCCAGAGTCCAGAGA 540
 DB 194 GlnProLysSerSerArgProProLeuSerGlnArgHisThrPheProGlnValGlnArg 213
 QY 541 ATGCCCAGTCAGATTAAGGACTTAAGTGGAGTCTTGAAGCAGAAAAGTTCT 600

DB 214 MetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnAlaGlnLysValPro 233
 QY 601 CATTAACGAGAGAGGCTGAATCACTATCTGTAGAAAACAAATTAACAAGATT 660
 DB 234 HisAsnGlnArgLysProGlnSerThrHisLeuLeuGlnAsnGlnAsnThrGlnGlnIle 253
 QY 661 CCAGCTGCCATTAGCAGTCTTCTCATGACAGCAAGCAACAGAGTGGCAAAACAGAT 720
 DB 254 ProLeuAlaIleSerSerSerSerPheThrThrSerHisSerValGlnAsnArgAsp 273
 QY 721 CATGAGAGAGCAGTACAGCCCTGTTCTCTCAGAGATCCAGCCCTCCAGCAGCTGCAGC 780
 DB 274 HisArgGlyGlyMetGlnProCysSerProGlnArgGlyGlnProAlaSerCysSer 293
 QY 781 CTCGACCAAAATATACCTCCCTATTAATTAACAAGCTGGAGACCACTTCCCAAAAG 840
 DB 294 ProHisGlnAsnIleLeuProTyrLysTyrThrSerTrpArgProProPheProLysArg 313
 QY 841 TCTGATGAAGAAGATGTCACAGACCAATGATGATGATGATGATGATGATGATGATGAT 900
 DB 314 SerAspArgLysAspValGlnHisAsnGlnTrpTyrIleGlyGlnTyrSerArgGlnAla 333
 QY 901 GTGGAGAGGCATTCATGAAGAGAGACAGAGATGATGATGATGATGATGATGATGATGAT 960
 DB 334 ValGlnGlnAlaPheMetLysGlnAsnLysAspLysSerPheLeuValArgAspCysSer 353
 QY 961 ACAAAATCCAAAGAGAGCCCTATGTTTGGCTGTGTTTATGAGCAAAAGTCTACAAT 1020
 DB 354 ThrLysSerLysGlnGlnProTyrValLeuAlaValPheTyrGlnAsnLysValTyrAsn 373
 QY 1021 GTAAAAATCCGCTCTCTGAGAGAGAAATCAGACAGTTGCTCCGAGACAGACTCAGAGA 1080
 DB 374 ValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlnTyrGlnLysLeuArgGly 393
 QY 1081 GATGAGAGATTGATTCAGTGAAGACATCATGACACTCAAGAAAT 1128
 DB 394 AspGlnLysPheAspSerValGlnAspIleIleGlnHisTyrLysAsn 409
 RESULT 4
 ID AAE22610 standard; Protein: 353 AA.
 XX
 AC AAE22610;
 XX
 DT 26-JUL-2002 (first entry)
 XX
 DE Human MIST splice variant protein from clone #12.
 XX
 KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
 KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
 OS Homo sapiens.
 XX
 PN MO200226986-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US30593.
 XX
 PR 29-SEP-2000; 2000US-237030P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
 DR WPI: 2002-372126/40.
 DR N-PSDB: AAD35802.
 XX
 PT New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma

XX Claim 11; Fig 8; 171pp; English.

PS The present invention relates to novel mast cell immunoreceptor signal

CC transducer (MIST) proteins and polynucleotides encoding such proteins.

XX MIST sequences of the invention are useful for preventing, treating or

CC ameliorating a medical condition in mammalian subject. They are useful

CC for treating an immune disorder involving hyperactivity of B- or T-

CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-

CC lymphoma, tumor or thymoma in a mammal. MIST sequences are useful

CC as targets for therapeutic intervention in immune cell disorders and

CC inflammatory indications, for diagnosis and/or screening of disorders

CC or diseases associated with expression of MIST, for screening for

CC antagonists or inhibitors of the interaction of MIST with cellular

CC signalling components. They are used in assays that detect activation

CC or induction of various B and T-cell-related neoplasms or cancers.

CC Sequences of the invention are also used in gene therapy. The present

CC sequence is human MIST splice variant protein from clone #12.

XX

SQ Sequence 353 AA:

Alignment Scores:

Pred. No.:	4,86e-163	Length:	353
Score:	1694.00	Matches:	317
Percent Similarity:	99.69%	Conservative:	1
Best Local Similarity:	99.37%	Mismatches:	1
Query Match:	83.86%	Indels:	0
DB:	23	Gaps:	0

US-09-856-061-3 (1-1129) x AAE22610 (1-353)

QY 172 ATGGAGAGACATGCGATGCTGATTAAATTTTACAGCCCGGCTATTAAGAATCGAA 231

DB 1 MetGluGluThrTrpLeuIleIleIleIleLeuProIleArgProIleIleGluSerGlu 20

QY 232 TATGCAATACACATATTTCAAGTTGCAATGACACTCCCTCCGTTAGACACAGG 291

DB 21 TyrAlaAspThrHisTyrPheIleValAlaMetAspThrProIleuAspThrArg 40

QY 292 ACCTGATTCCTCCATTTGGACGCGGACCTGGAACACACAGCGAGTTGGAAGATGGAC 351

DB 41 ThrSerIleSerIleGlyGlnProIleTrpAsnThrGlnThrArgLeuGlnAspValAsp 60

QY 352 AAACCATTTTCCAGGACGTCAGAACCAAAACATTTAAAGAGATGCATCCGTAAGAAG 411

DB 61 LysProIleSerIleLysAspValArgSerGlnAsnIleIleIleIleIleIleIleIleIle 80

QY 412 AACAGATTCCTTTACCACTCTCGGCTCTCATTAACACTCCGGAAGATCAACACC 471

DB 81 AsnLysIleProIleuProIleProIleArgProIleuIleThrLeuProIleLysTyrGlnPro 100

QY 472 TTGCCCCGAGCGGAGGACGACGACCTTATTCAGACACACTTCCGAA 531

DB 101 LeuProIleuProIleuProIleuSerArgProIleuSerIleuArgIleuIleuProIleu 120

QY 532 GTCCAGGAATGCCAGTACGATTAAGGACTTAAGTGAAGTCTTGAAGCAGAA 591

DB 121 ValGlnArgMetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnAlaGln 140

QY 592 AAAGTTCCTATTAACGAGGAAAGCTGATCACTGTTAGAAAACCAAAATACT 651

DB 141 LysValProHisAsnGlnArgLysProGlnSerThrHisLeuLeuGlnAsnThr 160

QY 652 CAAGAGATTCACATTCGCTTAGAGTCTTCATTCAGCAAGCAACAGCAGGTGCAA 711

DB 161 GlnGlnIleProLeuAlaIleSerSerSerPheThrThrSerAsnHisSerValGln 180

QY 712 AACAGATCATAGAGAGGACGACGACCTGTTCTCCAGAGATGCCAGCTCCAGCC 771

DB 181 AsnArgAspHisArgGlyGlyMetGlnProCysSerProGlnArgCysGlnProProAla 200

QY 772 AGCTGACGCCCTCAGCAAAATATATCTGCTATTAATACACAGCTGGAGACCACTTTC 831

DB 201 SerCysSerProHisGlnAsnIleLeuProTyrLysTyrThrSerTrpArgProPhe 220

QY 832 CCCAAAGGTCGTGATGAAGAAGATGTCACAGCAATGAATGATTCAGAAATACAGC 891

DB 221 ProLysArgSerAspArgLysAspValGlnHisAsnGluTrpTyrIleGlyGluTyrSer 240

QY 892 CGCAGGACGATGAGAGGACGATTCATGAAGAGACAGAGATGATGTTCTGTGTCGCA 951

DB 241 ArgGlnAlaValGlnGlnAlaPheMetLysGlnAsnLysAspLysPheLeuValArg 260

QY 952 GATTGTTCCACAATAATCCAGAGAGACCCATGTTTGCTGTGTTTATGGAACAA 1011

DB 261 AspCysSerThrLysSerLysGlnGluProTyrValLeuAlaValAPheTyrGlnAsnLys 280

QY 1012 GTCTACAATGTAAATATCCGCTTCGTGAGAGGAAATCAGCAGTTGCGCTGGGACAGCA 1071

DB 281 ValTyrAsnValLysIleArgPheLeuGlnArgAsnGlnGlnIleAlaLeuGlyThrGly 300

QY 1072 CTCAGAGAGATGAGAAATTTGATTGATGAGAGACATCATCGAACATCAAGAAAT 1128

DB 301 LeuArgGlyAspGlnLysPheAspSerValGlnAspIleIleGlnHisTyrLysAsn 319

RESULT 5

AAB81070

ID AAB81070 standard; Protein; 435 AA.

XX

AC AAB81070:

XX

XX 25-JUN-2001 (first entry)

DE Murine mast cell-specific signal transduction protein.

XX

KW Mast cell; signal transduction; mouse; allergic disease.

OS Mus musculus.

XX

PN JP3146204-B1.

XX

PD 12-MAR-2001.

XX

XX 17-SEP-1999; 99JP-0263778.

XX

PR 17-SEP-1999; 99JP-0263778.

XX

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX

DR WPI; 2001-310022/33.

DR N-PSDB; AAF86139.

XX

PT Mast cell-specific signal transduction molecule, useful for screening

PT therapeutic compounds for treating allergies, is specifically expressed

PT by mouse mast cell -

XX

PS Claim 1; Page 7-8; 12pp; Japanese.

XX

CC This sequence represents a murine mast cell-specific signal transduction

CC protein. The invention includes the cDNA and protein sequences of the

CC mast cell-specific signal transduction molecule and an expression vector

CC containing the polynucleotide sequence. The coding sequence of the signal

CC transduction protein can be used for screening therapeutic compounds

CC which will be useful for treating allergic diseases.

XX

SQ Sequence 435 AA:

Alignment Scores:

Pred. No.:	6.21e-106	Length:	435
Score:	1132.00	Matches:	229
Percent Similarity:	72.07%	Conservative:	42
Best Local Similarity:	60.90%	Mismatches:	101
Query Match:	56.04%	Indels:	4
DB:	22	Gaps:	3

US-09-856-061-3 (1-1129) x AAB81070 (1-435)

OY 121 GATGAGCAAAAGCCACATGATGACTATGATGACCCCTGCTGATGGAAG 180
 |||||
 Db 59 AspglyAlalysclYhisSerAspAsp*****AspAspProGluLeuArgMetGlu 78
 |||||
 OY 181 ACATGCGAGTCGATTA 198
 |||||
 Db 79 ThrTrpGlnSerIleLeu 84
 |||||
 RESULT 7
 AAM59866
 ID AAM59866 standard; Protein; 449 AA.
 AC AAM59866;
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE Amino acid sequence of the human BLNK-2 protein.
 XX
 KW Human: BLNK-2; B cells linker protein-2; apoptosis; Grb2; PLC-gamma;
 KW SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;
 KW ras pathway; GDP; GTP; calcium pathway; antigen.
 XX
 OS Homo sapiens.
 OS
 PN MO9832852-A1.
 PN
 PD 30-JUL-1998.
 PD
 PF 23-JAN-1998; 98WO-US01394.
 PF
 PR 17-MAR-1997; 97US-0819013.
 PR
 PR 24-JAN-1997; 97US-0788322.
 PR
 XX
 PA (UNITW) UNIV WASHINGTON.
 PA
 XX
 PI Chan AC, Fu C;
 PI
 DR WPI: 1998-427948/36.
 DR
 DR N-PSDB; AAV41902.
 DR
 XX
 PT Human B cell linker proteins - useful in the treatment of diseases
 PT Involving increased or decreased apoptosis
 PT
 XX
 PS Disclosure; Fig 3; 56pp; English.
 PS
 XX
 CC This is the amino acid sequence of the human BLNK-2 (B cells linker
 CC protein-2) protein, used the treatment of diseases involving the
 CC increase and decrease of apoptosis, in the method of the invention.
 CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
 CC intracellular calcium levels and Grb2 is an adapter molecule
 CC containing two SH3 domains that mediate its interaction with the
 CC guanine nucleotide exchange factor, Son of Sevenless (SOS) which in
 CC turn activates the ras pathway by facilitating the exchange of GDP for
 CC GTP on the ras molecule. Activation of both ras and calcium pathways
 CC are required for efficient B cell antigen receptor function. BLNK
 CC binding proteins can be used to identify BLNK proteins in a target
 CC sample.
 CC
 XX
 SQ Sequence 449 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1,78e-20 Length: 449
 Score: 290.50 Matches: 107
 Percent Similarity: 38.50% Conservative: 57
 Best Local Similarity: 25.12% Mismatches: 142
 Query Match: 14.38% Indels: 120
 DB: 19 Gaps: 18
 US-09-856-061-3 (1-1129) x AAM59866 (1-449)
 OY 60 CCACTACAGAGATGAACAAGCCTCTTCTAGACTGGAAAGAACTTT----- 108
 ||| |||:||||| ||| |||:||||| :|||:

Db 50 ProAlaAspGluGluGluGlnTrpSerAsp-AspHisSerAspTryGluAsnProAs 69
 OY 109 -----GCTCAGTCCGTGATGAGCAAAAGGCCACATGATGACTA 152
 |||:||||| :|||:||||| |||
 Db 69 pGluHisSerAspSerGluMetCtyrValMetProAlaGluAsnHisAspSerTyr 89
 |||:||||| :|||:||||| |||
 OY 153 TGATGACCCCTGACCTTCGGATGAGAGACATGCGATGATTAATTTTACAGCCG 212
 |||:||||| :|||:||||| |||
 Db 89 rGluProProValGluGluGln-----ThrArgProValHisProAlaLe 105
 |||:||||| :|||:||||| |||
 OY 213 GCCTATTAAGGAATCTGAATATGCATACACTATTTCAGTTGCAATGACACTCC 272
 |||:||||| :|||:||||| |||
 Db 105 uProPheAlaArgGlyGluTyr----- 112
 |||:||||| :|||:||||| |||
 OY 273 CCTTCGCTGACACCGACCTCTATCTCCATTGACACCGACCTGGAACACACAGAC 332
 |||:||||| :|||:||||| |||
 Db 113 -----IleAspAsnArgSerSer----- 118
 |||:||||| :|||:||||| |||
 OY 333 GAGGTGGAAGAGTGGCAAAACCATTTCCAGGACGTCAGAACCAAAATTAAGG 392
 |||:||||| :|||:||||| |||
 Db 119 -----GlnArgHisSerProPheSerIleThrLeuProSerTyr----- 132
 |||:||||| :|||:||||| |||
 OY 393 AGATGATCCGTAAAGAAAGAACAGATTCTTTACACCTCTCGGCTCTATACACT 452
 |||:||||| :|||:||||| |||
 Db 133 -----ProSerTrpProSerGluValAlaArgLeuThrSerThrLeuProAlaLeuThrAl 151
 |||:||||| :|||:||||| |||
 OY 453 TCCGAAGATACCAACCTTGCCTGAGCCG----- 486
 |||:||||| :|||:||||| |||
 Db 151 AleuGlnLysProGln--ValProProLysProLysGlyLeuLeuGlnAspGluAlaAs 170
 |||:||||| :|||:||||| |||
 OY 487 ----- 497
 |||:||||| :|||:||||| |||
 Db 170 pTyrValValProValGluAsnAspGluAsnTyrIleHisProThrGlnSerSer 190
 |||:||||| :|||:||||| |||
 OY 498 GCCA-----CCTTATCTCAGACACACCTTTCCAAAGTCCAG----- 537
 |||:||||| :|||:||||| |||
 Db 190 rProProGluLysAlaProMetValAsnArgSerThrLysProAsnSerSerThrPr 210
 |||:||||| :|||:||||| |||
 OY 538 -----GGAATGCCCGCATGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGG 593
 |||:||||| :|||:||||| |||
 Db 210 AlaSerProProGlyThrAlaSerGlyArgAsn--SerGlyAlaTrpGluThrLysSe 229
 |||:||||| :|||:||||| |||
 OY 594 A-----GTTCCTCATTAACAGAGAGCCT----- 618
 |||:||||| :|||:||||| |||
 Db 229 rProProProAlaAlaProSerProLeuProAlaArgAlaGlyLysLysProThrThrProLe 249
 |||:||||| :|||:||||| |||
 OY 619 -GAATCAATCATCTGTAGAAAACCAAAATACTCAAGAGATTCCACTTGCATTACAG 677
 |||:||||| :|||:||||| |||
 Db 249 uLysThrThrProValAlaSerGlnAsn----- 259
 |||:||||| :|||:||||| |||
 OY 678 TTCTTCATTTGACGACAAAGCAACACAGTGTGCAAAACAGAGATGATGAGAGAGC----- 732
 |||:||||| :|||:||||| |||
 Db 260 -AlaSerValLysGluGluGluLysProIleProAlaGluArgHisArgLysSerSerH 279
 |||:||||| :|||:||||| |||
 OY 733 -ATGACCCCTGTTCTCTCAGAGATGACAGCCCTCCAGCCAGTGCAGCCCTCAGAAA 791
 |||:||||| :|||:||||| |||
 Db 279 sArgGlnGlnAlaValGlnSerProValPheProProAlaGlnLysGlnIleHisGlnL 299
 |||:||||| :|||:||||| |||
 OY 792 TATACGCCCTAT--AAATACACA--AGCTGAGACCA-----CCTTTCC 833
 |||:||||| :|||:||||| |||
 Db 299 sProIleProLeuProArgPheThrGluGlyLysAsnProThrValAspGlyProLeuPr 319
 |||:||||| :|||:||||| |||
 OY 834 CAAGAAGTCGATGAAGAAGATGTCACAGCACATGA--TG 872
 |||:||||| :|||:||||| |||
 Db 319 oSerPheSerSerAsnSerThrIleSerGluGlnGlnAlaGlyValLeuCysLysProTr 339
 |||:||||| :|||:||||| |||
 OY 873 GTACATTGAGAAATACAGCCCGACGACGATGAAGAGGCATTCGAGAGAAACAAGA 932
 |||:||||| :|||:||||| |||
 Db 339 pTyrAlaGlyAlaCysAspArgLysSerAlaGluGlnAlaLeuHisArgSerAsnLysAs 359
 |||:||||| :|||:||||| |||
 OY 933 TGGTAGTTTCTTGGTCGAGATTGTTCCACAAATCCAGAGAAAGCCCTATGTTTGGC 992
 |||:||||| :|||:||||| |||
 Db 359 pGlySerPheLeuIleArgLysSerSerGlyHisAspSerLysGlnProTyrThrLeuVa 379


```

Db 286 sArgIngluaIaValGlnSerProValPheProProAlaGlnIleHisGlnIly 306
Oy 792 TATAGCCCTAT--AAATACACA--AGCTGGAGACCA-----CCTTCCC 833
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 sProIleProLeuProAlaGlnPheThrGlnGlnIleGlnIleProThrValAspGlyProLeuPr 326
Oy 834 CAAAGGCTGTGATGAAGAAGATGTCACGACATGAA-----TG 872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 oSerPheSerAsnSerThrIleSerGlnGlnIleGlnIleValLeuCysLysProTr 346
Oy 873 GTACATTTGAGAAATACAGCCGACAGTGGAGAGAGATTCATGAGAGACAAAGA 932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 pTyraIaGlnIleAlaCysAspArgIleGlnIleGlnIleAlaLeuHisArgSerIleLysAs 366
Oy 933 TGGTATTTCTTGTGTCGAGATTTGTCACAAATCCAGAAAGAGCCCTATGTTTGGC 992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 pGlySerPheLeuIleArgLysSerSerGlnHisAspSerLysGlnProTyThrLeuVa 386
Oy 993 TGTGTTTTTGAACAAGATGTCACATGTAATAATCCGCTCTGGAGAGAAATGACA 1052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 ValPhePheAsnLysArgValIleThrValIleProValIleArgPheIleGlnIleThrLysG 406
Oy 1053 GTTGGCTGTCGGAGAGACTCAGAGAGATGAGAAGTTTGATTCAGTACAGACATCAT 1112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 nTyraIaLeuGlnIleArgLysLysAsnGlnIleGlnIleGlnIleGlnIleIle 426
Oy 1113 CGAACACTACAGAAAT 1128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 eArgAsnHisGlnHis 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
AAW59865
ID AAW59865 standard; Protein; 456 AA.
AC
XX AAW59865;
XX
XX 20-NOV-1998 (first entry)
XX
XX Amino acid sequence of the human BLNK-1 protein.
XX
XX Human; BLNK-1; B cells linker protein-1; apoptosis; Grb2; PLC-gamma;
XX SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; Sos;
XX ras pathway; GDP; GTP; calcium pathway; antigen.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FT misc-difference 327 /note="encoded by AGC"
XX FT
XX PN WO9833852-A1.
XX
XX PD 30-JUL-1998.
XX
XX PF 23-JAN-1998; 98WO-0501394.
XX
XX PR 17-MAR-1997; 97US-0819013.
XX PR 24-JAN-1997; 97US-0788322.
XX
XX (UNIW ) UNIV WASHINGTON.
XX PA
XX PI Chan AC, Fu C;
XX DR WPI; 1998-427948/36.
XX DR N-PSDB; AAV41901.
XX
XX Human B cell linker proteins - useful in the treatment of diseases
XX PT involving increased or decreased apoptosis
XX
XX PS Claim 4; Fig 1; 56pp; English.
XX
XX This is the amino acid sequence of the human BLNK-1 (B cells linker
XX protein-1) protein, used the treatment of diseases involving the

```

```

CC increase and decrease of apoptosis. in the method of the invention.
CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
CC intracellular calcium levels and Grb2 is an adapter molecule
CC containing two SH3 domains that mediate its interaction with the
CC guanine nucleotide exchange factor, Son of Sevenless (Sos) which in
CC turn activates the ras pathway by facilitating the exchange of GDP for
CC GTP on the ras molecule. Activation of both ras and calcium pathways
CC are required for efficient B cell antigen receptor function. BLNK
CC binding proteins can be used to identify BLNK proteins in a target
CC sample.
XX
XX SQ Sequence 456 AA;
XX
XX Alignment Scores:
XX Pred. No.: 3,61e-20 Length: 456
XX Score: 287.50 Matches: 107
XX Percent Similarity: 38.50% Conservative: 57
XX Best Local Similarity: 25.12% Mismatches: 142
XX Query Match: 14.23% Indels: 120
XX DB: 19 Gaps: 18
XX
XX US-09-856-061-3 (1-1129) x AAW59865 (1-456)
Oy 60 CCAGTACAGAGAGATGAACAAGCCCTCTTACAGCTGGAAAGAACTT----- 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 ProIaAspGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 76
Oy 109 -----GCTCAGCTCTGATGAGCAAAAGCCACAGTATGATGACRA 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 pGlnHisSerAspSerGlnMetTyraIaMetProIaGlnIleGlnIleAsnLysAspSerTy 96
Oy 153 TGATGACCCCTGAGCTTGGATGAGAGACATGCGATGCAATTAATTTTACAGCCCG 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 rGlnProProProValGlnIleGlnIle-----ThrArgProValHisProIaIle 112
Oy 213 GCCTATTAAGAAATCTGAATATGACATACACATATTTCAAGTTGCAATGACACTCC 272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 uProPheAlaArgGlyGlnIleTyra----- 119
Oy 273 CCTCCGTTAGACACCGAGCCCTCTATCTCATTTGACACCGACCTGGAACACAGAC 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 -----IleAspAsnArgSerSer----- 125
Oy 333 GAGGTTGAAAGAGGTGACCAACCCATTTCAGAGAGCTGACAGCCAAACATTAAAG 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 -----GlnArgHisSerProPheSerLysThrLeuProSerLys----- 139
Oy 393 AGATGATCCGTAAAGAAAGAAAGATTCCTTTACACCTTCCTGCTCATACACT 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 -----ProSerTrpProSerGlnLysAlaArgLeuThrSerThrLeuProIaLeuThrAl 158
Oy 453 TCCGAAAGATACCAACCCCTTGGCCCGGAGCCG----- 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 AleuGlnLysProGln-----ValProProLysProLysGlnLeuLeuGlnAspGlnIaAs 177
Oy 487 ----- 497
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 pTyraValProValGlnAspAsnAspGlnAsnTyIleHisProThrGlnSerSer 197
Oy 498 GCCA-----CCTTATCTCAGAGACACACCTTTCAGAGTCCAG----- 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 rProProProGlnLysAlaPrometValAsnArgSerThrLysProAsnSerSerThrPr 217
Oy 538 -----GGAATGCCAGTCAGATGATTAAGGACCTTAAGCTGCTTGAAGAGAAA 593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 oAlaSerProProGlnThrAlaSerGlnArgAsn-----SerGlnAlaTrpGlnTrLysSe 236
Oy 594 A-----GTTCCATTAACAGAGAGAGCT----- 618
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 rProProProIaIaIaProSerProLeuProArgAlaGlnLysLysProThrTrProLe 256
Oy 619 -GAATCAACTCATCTGTAGAAAACAAATAACTCAAGAGATTCCACTTGCCATTAGCAG 677
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4: SEQ ID NO: 31602; 658bp + Sequence Listing: English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.
XX
XX Sequence 46 AA;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1.49e-15 Length: 46
XX Score: 238.00 Matches: 46
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 11.78% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-856-061-3 (1-1129) x AAM71296 (1-46)
XX
XX
XX 931 GATGGTAGTTCTGTCGAGATTGTCACAAATCCAGAGAGCCCTATGTTTG 990
XX |||||||
XX 1 AspGlySerPheLeuValArgAspCysSerThrLysSerLysGluIuProTyrValLeu 20
XX |||||||
XX 991 GCTGTGTTTATGAGAACAACTGTAATGTAATAATCCGCTTCTGGAGAGAAATCAG 1050
XX |||||||
XX 21 AlaValPheTyrGluAsnLysValTyrAsnValLysIleArgPheLeuGluIuArgAsnGln 40
XX |||||||
XX 1051 CAGTTGCCCTGGGGACA 1068
XX |||||||
XX 41 GlnPheAlaLeuGlyThr 46
XX
XX RESULT 15
XX AAM18973
XX ID AAM18973 standard; Protein: 46 AA.
XX
XX AC AAM18973;
XX
XX 12-OCT-2001 (first entry)
XX
XX DE Peptide #5407 encoded by probe for measuring cervical gene expression.
XX
XX KM Probe: human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
XX OS Homo sapiens.
XX
XX PN WO200157278-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00670.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI

XX WPI: 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 27: SEQ ID NO 23799; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENP: see A110068-A1128459). The present sequence is a peptide encoded
XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX can be used to produce a single exon microarray, which can be used for
XX measuring human gene expression in a sample derived from human cervical
XX epithelial cells. By measuring gene expression, the probes are therefore
XX useful in grading and/or staging of diseases of the cervix, notably
XX cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 46 AA;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1.49e-15 Length: 46
XX Score: 238.00 Matches: 46
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 11.78% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-856-061-3 (1-1129) x AAM18973 (1-46)
XX
XX
XX 931 GATGGTAGTTCTGTCGAGATTGTCACAAATCCAGAGAGCCCTATGTTTG 990
XX |||||||
XX 1 AspGlySerPheLeuValArgAspCysSerThrLysSerLysGluIuProTyrValLeu 20
XX |||||||
XX 991 GCTGTGTTTATGAGAACAACTGTAATGTAATAATCCGCTTCTGGAGAGAAATCAG 1050
XX |||||||
XX 21 AlaValPheTyrGluAsnLysValTyrAsnValLysIleArgPheLeuGluIuArgAsnGln 40
XX |||||||
XX 1051 CAGTTGCCCTGGGGACA 1068
XX |||||||
XX 41 GlnPheAlaLeuGlyThr 46
XX

Search completed: April 21, 2003, 12:20:53
Job time : 58.706 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 12:16:17 ; Search time 13.2707 Seconds

(without alignments)
5006.288 Million cell updates/sec

Title: US-09-856-061-3

Perfect score: 2020

Sequence: 1 ttccagacttcagctgcgc.....catgcacatacaagaatt.1129

Scoring table:

BLOSUM62		
Xgapop 10.0 , Xgapext 0.5		
Fgapop 10.0 , Fgapext 0.5		
Delop 6.0 , Delext 7.0		

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p model -DEV=xlp
-Q/cgcn2.1/USPTO_SPOOL/US09856061/runat.21042003.113020.3774/app_query.fasta.1.3150
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061.ecgn1.1.1.46.0runat.21042003.113020.3774 -NCP=6 -ICP=3
-NO_XIPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgcn2.6/ptodata/1/1aa/5A.COMB.pep.*
- 2: /cgcn2.6/ptodata/1/1aa/5B.COMB.pep.*
- 3: /cgcn2.6/ptodata/1/1aa/6A.COMB.pep.*
- 4: /cgcn2.6/ptodata/1/1aa/6B.COMB.pep.*
- 5: /cgcn2.6/ptodata/1/1aa/PTUS.COMB.pep.*
- 6: /cgcn2.6/ptodata/1/1aa/Backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	287.5	14.2	456	2	US-08-819-013-1
2	155	7.7	724	1	US-07-906-349A-5
3	155	7.7	724	1	US-08-167-035-2
4	155	7.7	724	1	US-08-208-887A-2
5	155	7.7	724	2	US-08-539-005-2
6	150.5	7.5	724	4	US-09-280-598-5
7	144.5	7.2	722	4	US-08-390-874C-12
8	144.5	7.2	722	4	US-09-265-772-12
9	143	7.1	659	1	US-08-391-615-2
10	140	6.9	659	4	US-08-426-509A-8
11	140	6.9	659	4	PCT-US95-05008-8
12	137	6.8	442	1	US-08-391-615-6

13	130	6.4	92	1	US-08-202-389-28	Sequence 28, App1
14	129.5	6.4	474	3	US-08-729-416C-1	Sequence 1, App1
15	129.5	6.4	594	3	US-08-729-416C-7	Sequence 7, App1
16	125	6.2	593	1	US-08-202-389-12	Sequence 12, App1
17	125	6.2	593	1	US-08-018-129-5	Sequence 5, App1
18	125	6.2	593	2	US-08-448-250-5	Sequence 5, App1
19	124.5	6.2	160	1	US-08-167-035-40	Sequence 40, App1
20	124.5	6.2	160	1	US-08-208-887A-40	Sequence 40, App1
21	124.5	6.2	160	2	US-08-539-005-40	Sequence 40, App1
22	124.5	6.2	160	4	US-09-280-598-37	Sequence 37, App1
23	123	6.1	454	4	US-08-259-264-2	Sequence 2, App1
24	121.5	6.0	1893	1	US-08-188-582-11	Sequence 11, App1
25	121.5	6.0	1893	1	US-08-646-715-11	Sequence 11, App1
26	120	5.9	694	3	US-08-559-397A-31	Sequence 31, App1
27	119	5.9	99	2	US-08-820-754-25	Sequence 25, App1
28	119	5.9	99	3	US-08-956-652-25	Sequence 25, App1
29	119	5.9	99	3	US-08-956-869-25	Sequence 25, App1
30	119	5.9	99	3	US-08-948-547-25	Sequence 25, App1
31	119	5.9	620	4	US-08-426-509A-9	Sequence 9, App1
32	119	5.9	620	5	PCT-US95-05008-9	Sequence 9, App1
33	119	5.9	1872	1	US-08-188-582-14	Sequence 14, App1
34	119	5.9	1872	1	US-08-646-715-14	Sequence 14, App1
35	117.5	5.8	474	3	US-08-807-342B-2	Sequence 2, App1
36	117	5.8	95	1	US-08-308-086-8	Sequence 8, App1
37	117	5.8	96	2	US-08-479-078-20	Sequence 20, App1
38	117	5.8	102	1	US-08-202-389-27	Sequence 27, App1
39	116	5.7	1130	4	US-09-442-100-4	Sequence 4, App1
40	115	5.7	675	4	US-08-426-509A-4	Sequence 4, App1
41	115	5.7	675	5	PCT-US95-05008-4	Sequence 4, App1
42	114.5	5.7	625	1	US-08-391-615-3	Sequence 3, App1
43	113	5.6	226	3	US-09-195-868-30	Sequence 30, App1
44	113	5.6	1229	3	US-09-195-868-28	Sequence 28, App1
45	112	5.5	99	3	US-09-195-868-23	Sequence 23, App1

ALIGNMENTS

RESULT 1
US-08-819-013-1
Sequence 1, Application US/08819013
Patent No. 5994522
GENERAL INFORMATION:
APPLICANT: Chan, Andrew C.
TITLE OF INVENTION: BLNK PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,013
FILING DATE: 17-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/788,322
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-64383-1/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:

```

: SEQUENCE CHARACTERISTICS:
:   LENGTH: 456 amino acids
:   TYPE: amino acid
:   STRANDEDNESS: unknown
:   TOPOLOGY: unknown
:   MOLECULE TYPE: protein
:   US-08-819-013-1

Alignment Scores:
Pred. No.:      2,72e-21      Length:      456
Score:          287.50        Matches:      107
Percent Similarity: 38.50%    Conservative: 57
Best Local Similarity: 25.12%  Mismatches:   142
Query Match:     14.23%      Indels:      120
DB:              2           Gaps:          18

US-09-856-061-3 (1-1129) x US-08-819-013-1 (1-456)

QY 60 CCAGTACGAGAGATGAACAGCCCTCTTCTAGACTGGGAAAGAACTTT----- 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 ProAlaAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu
QY 109 -----GCTGCAGTCCCTGGAGTGAAGCAAAAGCCACAGTATGATGACTA 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 pEluHisSerAspSerGluMetLysValMetProAlaGluGluGluGluGluGluGlu
QY 153 TGATGACCCCTGACCTTGGGATGAGAGACATGCGAGTCGATTAATTTTACCGCCCG 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 rGluProProProAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGlu
QY 213 GCCTATTAAGAAATCTGATATAGATACACACTATTTTCAAGTTCGAATGACACTCC 272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 uProPheAlaArgGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLys
QY 273 CCTTCCGTTAGACACAGACCTCTATCTCCATTGGACACCGGACCTGGAACACAGAC 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 -----IleAspAsnArgSerSer----- 125
QY 333 GAGGTTGGAAGAGTGGCAAAACCCATTTCAGAGCGTGCAGAGCCAAACATTAAGG 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 -----GlnArgHisSerProProPheSerLysThrLeuProSerLys----- 139
QY 333 AGATGATCTCGCTAAGAAAGAACAGATTCCCTTACCACCTCTCGCCCTCATTAACACT 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 -----ProSerTrpProSerGluLysAlaArgLeuThrSerThrLeuProAlaLeuThrAl 158
QY 453 TCCGAGAGATGACCAACCTTGCCCGCCCTGAGCCG----- 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 aLeuGluLysProGln-----ValProProLysProLysGlyLeuLeuGluAspGluAlaAs 177
QY 487 -----GAGAGCAGCAG 497
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 pTyrValValProValGluAspAsnAspGluAsnTyrIleHisProThrIleGluSerSerSe 197
QY 498 GCCA-----CCTTATCTCAGAGACACACCTTTCAGAGTCCAG----- 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 rProProProGluLysAlaProMetValAsnArgSerThrLysProAsnSerSerThrPr 217
QY 538 -----GGAATGCCAGTGCATAGACTTAAGGACTTAAGTGAAGCTCTTGAAGCAAGAA 593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 oAlaSerProGluThrAlaSerGlyArgAsn-----SerLysAlaTrpGluThrLysSe 236
QY 594 A-----GTTCCTCATTAACCAAGAGAGAGAGCT----- 618
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 rProProProAlaAlaProSerProLeuProArgAlaGlyLysLysProThrThrProLe 256
QY 619 -GAATCAACTATCTGTGTAAGAAACCAAAATACTCAAGAGATTCACCTTGCATTACGACG 677
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 uLysThrThrProValAlaSerGlnAsn----- 266
QY 678 TCTTCATTCACGACCAACCAACAGTGTGCACAAACAGACATCATAGAGAGCG----- 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 -AlaSerSerValGlyGluGluLysProIleProAlaGluArgHisArgGlySerSerHi 286

```

```

QY 733 -ATGAGCCCTGTCTCTCCACAGATGCCAGCCTCCAGCCACTCCAGCCCTCAGGAAA 791
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 sArgGlnGluAlaValGlnSerProValPheProProAlaGlnLysGlnIleHisGlnLys 306
QY 792 TATACTGCCCTTAT-----AAATACACA---AGCTGAGACCA-----CCTTCC 833
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 sProIleProLeuProAlaGlyPheThrGluGlyGlyLysProThrValAspGlyProLeuPr 326
QY 834 CAAAAGTCTGTGATGAAAGATGTCCACACATGAA-----TG 872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 oIlePheSerSerAsnSerThrIleSerGluGlnGluAlaGlyValLeuGlyLysProTr 346
QY 873 GTACATTTGAGATACACAGCCGCCAGCAGTGCAGAGCATTCATCAGACGACAGACAGA 932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 pTyrAlaGlyAlaCysAspArgLysSerAlaGluAlaLeuHisArgSerAsnLysAs 366
QY 933 TGTGATTTCTTGTGTCGAGATTTGTCCAAAAATCCAAAGAGAGCCCTATGTTTGAGC 992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 pGlySerPheLeuIleArgLysSerSerGlyHisAspSerLysGlnProTyrThrLeuVa 386
QY 993 TGTGTTTATGAGCAAAAGTCTACATGTAAATCCGCTTCTGAGAGAGATCACA 1052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 lValPhePheAsnLysArgValTyrAsnIleProValArgPheIleGluAlaThrLysG 406
QY 1053 GTTTCCTCGGGACAGACACTCAGAGAGATGAGAAAGTTTGATTCAGTAGAAGACATCAT 1112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 nTyrAlaLeuGlyArgLysLysAsnGlyLysGluLysPheGlySerValAlaGluIle 426
QY 1113 CGAACACTACAGAAT 1128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 eArgAsnHisGlnHis 431

```

```

RESULT 2
US-07-906-349A-5
: Sequence 5, Application US/07906349A
: Patent No. 5434064
:
: GENERAL INFORMATION:
: APPLICANT: Schlessinger, Joseph
: APPLICANT: Skolnik, Edward Y.
: APPLICANT: Margolis, Benjamin I.
: TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
: TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES
: NUMBER OF SEQUENCES: 16
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Browdy and Neimark
: STREET: 419 Seventh Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/906,349A
: FILING DATE: 30-JUN-1992
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/643,237
: FILING DATE: 18-JAN-1991
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 724 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear

```

MOLECULE TYPE: protein
US-07-906-349A-5

Alignment Scores:

Pred. No.: 1.79e-07 Length: 724
Score: 155.00 Matches: 94
Percent Similarity: 37.24% Conservative: 52
Best Local Similarity: 23.98% Mismatches: 103
Query Match: 7.67% Indels: 143
DB: 1 Gaps: 20

US-09-856-061-3 (1-1129) x US-07-906-349A-5 (1-724)

```

OY 308 GACAGCCAGCTGGAACACACAGAGAGTGAAGTGAACAAC----- 355
DB 47 GIUAIAlaIqPro-----ArgHrAsnGlyTrpLeuAsnGlyTrpAsnGluThrGly 64
OY 356 -----CCATTTCAGGAGCTCAGAACCAATTAAGAGAGATGCTGCTAA 406
DB 65 GIUysGlyAspPheProGlyThr-----TyrValGluTyrIleGly-Argly 80
OY 407 GAAAGAACAAAGTTCTTACACCTCTCGGCTCTACACTTCGGAAGATACC 466
DB 80 slyIleSerPro-ProThrProlysProAlqPro-----ProArg----- 93
OY 467 AACCTTGCCCGCCGAGCGAGAGCAGCAGG----- 498
DB 94 --ProleuProValAlaProGlySerSerLyThGluAlaAspValGluGlnAla 113
OY 499 -----CCACCTTATCTCAGAGACACCTTCCAGAGTC----- 534
DB 113 euThrLeuProAspLeuAlaGluGlnPheAlaProProAspIleAlaProProLeuLeu 133
OY 535 -----CAGGAATG-----C 544
DB 133 IeLyLeuValGluAlaIleGluLyLyGlyLeuGluCysSerThrLeuTyrArgThG 153
OY 545 CCAGTCAGATTAAGTAAAGGACTTAAGAGCTGCTGAGCAGACAGAAAAGTTCTCAT 604
DB 153 InSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 172
OY 605 ACCAGAGAGAGCTGAGTCACTCATCTGTAGAAAACCA----- 645
DB 172 alAspLeuIleMetIleAspValHisValLeuAlaAspAlaPheLyArgTyrLeuLeu 192
OY 646 -----AATCTCAAGAGATTCACCTGCATTAAGCAGTCT----- 681
DB 192 spleuProAsnProValIleProAlaAlaValTyrSerGluMetIleSerLeuAlaProG 212
OY 681 ----- 681
DB 212 IuValGInSerSerGluGluTyrIleGlnLeuLeuLyLyLeuIleArgSerProSerI 232
OY 682 -----TCAT 685
DB 232 IeProHISGlnTyrTrpLeuThrLeuGlnTyrLeuLeuLySHISpHePheLyLeuSerG 252
OY 686 TCACGCAAGCAACACAGCTGTCGCAAAACAGAGATCATAGAGGACATGCAAGCTGTT 745
DB 252 InThrSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 271
OY 746 CTCCTCAGA-----GATGCAAGCTCCAGCCAGCTGACGAGCTCAGGAAA---ATATAC 796
DB 272 LeuPheArgPheSerAlaAlaSerSerAspSerThrGluAsnLeuIleLyValIleGlu 291
OY 797 TGCCCTATAATACACAGCTGAGA-----CCACCTTTC 832
DB 292 IleLeuIleSerThrGlu--TrpAsnGluArgGlnProAlaProAlaLeuProProLys 311
OY 833 CCAAAAG-----TCTGATACAAAGAGATGTCAGACACA 865
DB 311 roProLysProThrThrValAlaAsnAsnGlyMetAsnAsnMetSerLeuGlnAsnA 331

```

```

OY 866 ATGATGCTACATTGGAGATATACCCCGCAGAGCAGTGAAGAGCATTCATGAAGAGA 925
DB 331 IaGluTrpTyrTrpGlyLysPheSerArgGluGlnValAsnGluLyS---LeuArgAspT 350
OY 926 ACAAGATGATGATTTCTTGGTCGAGATTTGTCACAAATCCAGAGACAGCCCTATG 985
DB 350 hrAlaAspGlyThrPheLeuValArgAspAlaSerThrLySerMetHisGlyAsp---TyrT 369
OY 986 TTTGGCTGTGTTTATATAGAACAAAGCTCAATGTAAA---ATCCGCTTCCTGAGAA 1042
DB 369 hrLeuThrLeu-----ArgLyGlyGlyAsnAsnLyLeuIleLySlePheHisA 386
OY 1043 GGAATCAGACAGTTTCCCTCGGAGCAGACTCAGAGAGATGAGAAGTTGATTCAGTAG 1102
DB 386 rGAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheSerSerVal 402
OY 1103 AAGACATCATCGACACTRCAAGAAAT 1128
DB 402 alGluLeuIleAsnHisTyrArgAsn 410

```

RESULT 3

US-08-167-035-2

Sequence 2, Application US/08167035

Patent No. 5618691

GENERAL INFORMATION:

APPLICANT: Schlessinger, Joseph

APPLICANT: Skolnick, Edward Y.

APPLICANT: Margolis, Benjamin L.

TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR

TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE

TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSER: PENNIE & EDMONDS

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: 10036-2711

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/167,035

FILING DATE: 16-DEC-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7683-062

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 724 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-167-035-2

Alignment Scores:

Pred. No.: 1.79e-07 Length: 724

Score: 155.00 Matches: 94

Percent Similarity: 37.24% Conservative: 52

Best Local Similarity: 23.98% Mismatches: 103

Query Match: 7.67% Indels: 143

DB: 1 Gaps: 20

US-09-856-061-3 (1-1129) x US-08-167-035-2 (1-724)

```

QY 308 GACAGCCGACCTGGAACACACAGAGGTTGGAAGAGTGACAAC----- 355
Db 47 GLUALAARPro-----ArgArgAsnGlyTTPLeuAsnGlyTYrAsnGluThrThrGly 64
QY 356 -----CCATTTCAGGAGCGTCAGAACCCAAACATTAAAGAGATGATCCGTAA 406
Db 65 GLuysGlyAspPheProGlyThr-----TyValGluTyrIleGly-Argly 80
QY 407 GAAAGAACAGATTCCTTACCACTCGGCTCTCATATCACTTCCGAAGAGTACC 466
Db 80 slyslIleSerPro-ProThrProLysProArgPro-----ProArg----- 93
QY 467 AACCTTGCCTCCGAGCGAGACAGCAG----- 498
Db 94 --ProLeuProValAlaProGlySerSerLysThGluAlaAspValGluGlnAlaL 113
QY 499 -----CCACCTTATCTAGAGACACACCTTTCAGAAAGTC----- 534
Db 113 euThrLeuProAspLeuAlaGluGlnPheAlaProProAspLeuAlaProProLeuLeu 133
QY 535 -----CAGGCAATG-----C 544
Db 133 IeLysLeuValGluAlaIleGluLysLysGlyLeuGluCysSerThrLeuTyrArgThrG 153
QY 545 CCAGTCAGATAGCTTAAAGGACTTAAGTGAAGTCTTGAAGCAGAAAGTCTCATTA 604
Db 153 InSerSerAsnLeuAlaGluLeuArgGlnLeuLeuAspCysAsp---ThrProSerV 172
QY 605 ACCAGAGAGACCTGATCACTATCTGTTAGAAACCA----- 645
Db 172 AlaAspLeuIleMetIleAspValIleValLeuAlaAspAlaPheLysArgTyrLeuLeuA 192
QY 646 -----AATACTCAAGAGATTCCTTCCACTTGCCTTAGCAGTTCT----- 681
Db 192 spleuProAsnProValIleProAlaAlaValTyrSerGluMetIleSerLeuAlaProG 212
QY 681 ----- 681
Db 212 IuValGlnSerSerGluGluTyrTyrIleGlnLeuLysLeuLysLeuLeuArgSerProSerI 232
QY 682 -----TCAT 685
Db 232 IeProHisGlnTyrTyrLeuThrLeuGlnTyrLeuLeuLysHisPhePheLysLeuSerG 252
QY 686 TCAGAGACAGAACCCAGAGTGTGCAAAAACAGAGATCATAGAGAGCATGAGCCCTGT 745
Db 252 InThrSerSerLysAsnLeuLeuAsnAlaArg-ValLeuSerGluIlePheSerProMet 271
QY 746 CTCCTCAGA-----GATGCCAGCCTCCAGCAGCAGCTGAGCCCTCAGGAAA---ATATAC 796
Db 272 LeuPheArgPheSerAlaAlaSerSerAspAsnThrGlnLeuIleLysValIleGlu 291
QY 797 TGCCCTAATATACACAGCTGAGAG-----CCACCTTCC 832
Db 292 IleLeuIleSerThrGlu--TyrAsnGlnArgGlnProAlaProAlaLeuProProLysP 311
QY 833 CCAAAAGC-----TCTGATGAAGAGATGCTCCAGACA 865
Db 311 ropProLysProThrThrValAlaAsnAsnGlyMetAsnAsnAsnMetSerLeuGlnAsnA 331
QY 866 ATGAATGTGTCATGTGAGATATACGCCGCCAGAGATGGAAGAGCATTCATGAGAGAGA 925
Db 331 IeGluTyrTyrTyrGlyAspIleSerArgGlnGluValaGlnGluLys---LeuArgAsp 350
QY 926 ACAAGATGATGATTTCTTGTGCGAGATGTTCCACAAATCCAAAGAGAGAGCCCTATG 985
Db 350 hrAlaAspGlyThrPheLeuValaIArgAspAlaSerThrLysMetHisGlyAsp---TyrT 369
QY 986 TTTTGGCTGTGTTTATGAGAACAAAGCTACAAATGTAA---ATCCGCTTCTCGAGAGA 1042
Db 369 hrLeuThrLeu-----ArgLysGlyLysAsnLeuLysLeuIleLysIlePheHisA 386

```

```

QY 1043 GGAATCAGCAGTTTCCTCGGGGACAGAGCTCAGAGAGATGAGAGATTTCATTCAGTAG 1102
Db 386 rGAspGlyLysTyrIlePheSerAspProLeu-----ThrPheSerValV 402
QY 1103 AAGACATCATCGAACACATCAAGAGT 1128
Db 402 alGluLeuIleAsnHisTyrArgAsn 410

RESULT 4
US-08-208-887A-2
: Sequence 2, Application US/08208887A
: Patent No. 5677421
: GENERAL INFORMATION:
: APPLICANT: Schlessinger, Joseph
: APPLICANT: Skolnick, Edward Y.
: APPLICANT: Margolis, Benjamin L.
: TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
: TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
: TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
: NUMBER OF SEQUENCES: 51
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PENNIE & EDMONDS
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: 10036-2711
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/208,887A
: FILING DATE: 11-MAR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Cornuzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7683-063
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SPO ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 724 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-208-887A-2

Alignment Scores:
Pred. No.: 1.79e-07 Length: 724
Score: 155.00 Matches: 94
Percent Similarity: 37.24% Conservative: 52
Best Local Similarity: 23.98% Mismatches: 103
Query Match: 7.67% Indels: 143
DB: 1 Gaps: 20

US-09-856-061-3 (1-1129) x US-08-208-887A-2 (1-724)
QY 308 GACAGCCGACCTGGAACACACAGAGGTTGGAAGAGTGACAAC----- 355
Db 47 GLUALAARPro-----ArgArgAsnGlyTTPLeuAsnGlyTYrAsnGluThrThrGly 64
QY 356 -----CCATTTCAGGAGCGTCAGAACCCAAACATTAAAGAGATGATCCGTAA 406
Db 65 GLuysGlyAspPheProGlyThr-----TyValGluTyrIleGly-Argly 80
QY 407 GAAAGAACAGATTCCTTACCACTCGGCTCTCATATCACTTCCGAAGAGTACC 466
Db 80 slyslIleSerPro-ProThrProLysProArgPro-----ProArg----- 93

```

[illegible]

```

GENERAL INFORMATION:
APPLICANT: Schlensing, Joseph
APPLICANT: Skolnick, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: 10036-2711
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,005
FILING DATE: 4-OCT-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/167,035
FILING DATE: 16-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-539-005-2

Alignment Scores:
Pred. No.: 1.79e-07 Length: 724
Score: 155.00 Matches: 94
Percent Similarity: 37.24% Conservative: 52
Best Local Similarity: 23.98% Mismatches: 103
Query Match: 7.67% Indels: 143
DB: 2 Gaps: 20

US-09-856-061-3 (1-1129) x US-08-539-005-2 (1-724)
QY 308 GACAGCCGACCTGGAGACACAGAGCGAGTGGAAAGGTGACAAAC----- 355
DB :::::::::::::::::::: ||| ||||| |||
47 GUAlAaArgPro-----ArgArGAsnGlyTrPrLeuAsnGlyTYrAsnGluThrThrcly 64
356 -----CCATTTCAGAGCGAGCGAGAGCCAAACATTAAAGGAAGACATCCGCTAA 406
DB ||||||| ||||||| ||| :::::::::::::::::::: :|||
65 GluLYsGlyAspPheProGlyThr-----TYrValGluTYrIleGly-Argly 80
QY 407 GAAGAGAACAGATTCTTACACACCTCTCGGCGCTCTCATTAACACTTCCGAAGATACC 466
DB ||||| ||| ||| ||||||| ||| :::::::::::::::::::: :|||
80 slystIleSerPro-PrOThrPrOlySProArgPro-----ProArg----- 93
QY 467 AACCTTGCCCGCTAGCCGAGAGACAGAG----- 498
DB 94 --ProLeuProValAlaAlaProGlySerSerLYsThrGluAlaAspValGluGlnGlnAlaL 113
QY 499 -----CCACCTTATCTAGAGACACACACTTCCGAAGATC----- 534
DB 113 eutHrLeuProAspLeuAlaGluGlnGlnPheAlaProProAspIleAlaProProLeuLeuI 133

```



```

QY 681 ----- 681
Db 212 GluValGlnSerSerGluGluTyrIleGlnLeuLeuLysLysLeuIleArgSerProSer 231
QY 682 ----- TCA 684
Db 232 IleProHisGlnTyrTyrLeuThrLeuGlnTyrLeuLeuLysHisPhePheLysLeuSer 251
QY 685 TTCAGCAGCAGCAGCAGCAGTGTGCAAAACAGAGATCATAGAGAGGCGCATGACCCCTGT 744
Db 252 GlnThrSerSerLysAsnLeuLeuAsnAlaArg-ValLeuSerGluIlePheSerProSe 271
QY 745 TCTCCACACA-----GATGCCAGCCTCCAGCCAGCTGACGCCCTCAGCAAA--ATATA 795
Db 271 LeuPheArgPheSerAlaIleSerSerAspAsnThrGluAsnLeuIleLysValIleG1 291
QY 796 CTGCCTATATAATACACAGCTGAGACA-----CCACCTTTC 831
Db 291 uIleLeuIleSerThrGlu--TTPAsnGluArgGlnProAlaProAlaLeuProProlys 310
QY 832 CCCAAAAG-----TCTGATAGAAAAGATGTCCAGCAC 864
Db 311 ProProlysProThrThrValAlaAsnAsnGlyMetAsnAsnMetSerLeuGlnAsn 330
QY 865 AATGATGTGACATTTGGAGAAATACAGCCCGCAGCAGCAGTGAAGAGCGCATTCAGAGAG 924
Db 331 AlaGluTyrPyrTrpIleAspIleSerArgGluGluValAsnGluLys--LeuArgAsp 349
QY 925 AACAGAGATGATGTTCTTCTGTCGCGAGATTTGCCAATAATCCAAAGAAAGACCCCTAT 984
Db 350 ThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp--Tyr 368
QY 985 GTTTGGCTGTGTTTATAGAAACAAGTCTACAATGTAAA--ATCCGCTTCCGTGAG 1041
Db 369 ThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIlePheHis 385
QY 1042 AGGAATCAGCAGTTGGCCTGGGAGACAGCAGCAGTGAAGAGATGAGAAATTGATTGCTGA 1101
Db 386 ArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheSerVal 401
QY 1102 GAAGACATCATGCACACTACAGAAAT 1128
Db 402 ValGluLeuIleAsnHisTyrArgAsn 410

RESULT 7
US-08-390-874C-12
; Sequence 12, Application US/08390874C
; Patent No. 6043062
; GENERAL INFORMATION:
; APPLICANT: Kimpel, Anke
; TITLE OF INVENTION: A Constitutively Active
; TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,874C
; FILING DATE: 17-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.

```

```

; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 02307K-057000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-390-874C-12

Alignment Scores:
Pred. No.: 2,19e-06 Length: 722
Score: 144.50 Matches: 75
Percent Similarity: 32.76% Conservative: 40
Best Local Similarity: 21.37% Mismatches: 95
Query Match: 7,15% Indels: 141
DB: Gaps: 17

US-09-856-061-3 (1-1129) x US-08-390-874C-12 (1-722)
QY 421 CCTTACCACTCTCGGCTCTCATACACTTCGAGAGATGCCAGTATACCTTGCCCTT 480
Db 85 ProThrProlysProArgPro-----ProArg-----ProLeuProVal 97
QY 481 GAGCCGAGAGCAGCAGG-----CCACCT 504
Db 98 AlaProGlySerSerLysThrGluAlaAspThrGluGlnGlnAlaLeuProLeuProAsp 117
QY 505 TTATTCAGAGACACACACTTTCAGAGATGCCAGGAAATGCCAGTACATACCTTAAG 564
Db 118 LeuAlaGluGlnPheAlaProProAspVal---AlaProProLeuLeuIleLysLeu 136
QY 565 GACTTAAGTGAGGTC-----CTTGAAGCAGAAAAGTTCTCATACCAAGCAGAGACCT 618
Db 137 GluAlaIleGluLysLysGlyLeuGluLysSerThrLeuLysArgThrGlnSerSer 156
QY 619 GAATCAACTATCTGTAGAA-----AACCAAAATACTCAAG 657
Db 157 AspProAlaGluLeuArgGlnLeuLeuAspCysAspAlaAlaSerValaAspLeuGluMet 176
QY 658 ATTCCACTGCATTTACAGCTTCTCATTCACGACGACGACCAACAGCAGTGTCAAGAGA 717
Db 177 IleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuAlaAspLeuProAsn 195
QY 718 GATCATAGAGAGGAGCATGACCCCTGTTCTCT----- 750
Db 196 -----ProValIleProValAlaValaTyrAsnGluMetSer 208
QY 751 -----CAGAGATGCCAGCTCCAGCCAGCTGC----- 777
Db 209 LeuAlaGlnGluLeuGlnSerProGluAspCysIleGlnLeuLeuLysLysLeuIleArg 228
QY 778 -----AGCCCTCAGAAAATATATAGTCCCTATATAATAC----- 810
Db 229 LeuProAsnIleProHisGlnGlyStrPheuThrLeuIleuLysHisPhePhe 248
QY 810 ----- 810
Db 249 LysLeuSerGlnAlaSerSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePhe 268
QY 810 ----- 810
Db 269 SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGluHisLeuLys 288
QY 811 -----ACAAGCTGG----- 819
Db 289 AlaIleGluLeuLeuIleSerThrGluTyrPsnGluArgGlnProAlaProAlaLeuPro 308
QY 820 ---AGACCACCTTTCCCC-----AAAAGCTGTAGTAAAGATGTC 858

```

```

Db      309  ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328
      859  CAGCACAATGATGTGATCTTGGAGAAATACAGCCCGCAGCAGTGAAGAAGCATTCATG 918
      329  GlnAspAlaGlnIuTrpIuTrpGlyAspIleSerArgIuValAsnMetLys-----Leu 346
      919  AAGGAGACAAGAGATGATGTTCTGTCGCGATGTTCCACAAATCCAAAGAGAG 978
      347  ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrIysMetHisGlyAsp 366
      979  CCTATGTTTGGCTGTGTTTATGAGAACAAAGTCTACAAATGTAA--ATCCGCTTC 1035
      367  ---TyrThrLeuThrLeu-----ArgIysGlyGlyAsnAsnLysLeuIleIysIle 382
      1036  CTGGAGAGAAATCAGCAGTTGGCCCTGGGACAGCAGCTCAGAGAGATGAGAACTTGAT 1095
      383  PheHisArgAspGlyLysIuTrpGlyPheSerAspProLeu-----ThrPheAsn 398
      1096  TCAGTAGAAGACATCATCGACATCAGACATACAGAAAT 1128
      399  SerValIuAlaGluLeuIleAsnHisTyrArgAsn 409

RESULT 8
US-09-265-772-12
Sequence 12, Application US/09265772
Patent No. 6300111
GENERAL INFORMATION:
APPLICANT: Klippel, Anke
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: A Constitutively Active
TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,772
FILING DATE: 10-MAR-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,874
FILING DATE: 17-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 02307K-057010US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 722 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-265-772-12

Alignment Scores:
Pred. No.: 2.19e-06
Score: 144.50
Percent Similarity: 32.76%
Best Local Similarity: 21.37%
Length: 722
Matches: 75
Conservative: 40
Mismatch: 95

```

```

Query Match: 7.15% Indels: 141
DB: 4 Gaps: 17
US-09-856-061-3 (1-1129) x US-09-265-772-12 (1-722)
      421  CCTTACACACCTCCGCGCTCTCATTAACACTTCCGAGCAATACCAACCTTCGCCCT 480
      85  ProThrProLysProArgPro-----ProArg-----ProLeuProVal 97
      481  GAGCCGAGAGCAGCAGC-----CCACCT 504
      98  AlaProGlySerSerLysThrGluAlaAspThrGluGlnAlaLeuProLeuProAsp 117
      505  TTAATCTCAGACACACACCTTCCAGAAATCCAGATCCAGATTAAGTAAAG 564
      118  LeuAlaGluGlnPheAlaProProAspVal---AlaProProLeuLeuIleLysLeuLeu 136
      565  GACTTAAGTGAGTGC-----CTTGAGCAGCAAAAAGTTCCTCATTAACCAAGAGAGCT 618
      137  GluAlaIleGluLysLysGlyLeuGluLysSerThrLeuTyrArgThrGlnSerSer 156
      619  GAATCACTCATCTGTAGAA-----AACCAAAATACTCAAGAG 657
      157  AsnProAlaGluLeuArgGlnLeuLeuAspCysAspAlaAlaSerValAspLeuGluMet 176
      658  ATTCCACTTGCATAGCAGCTTCTCATTCACAGACAGACACCAACCAAGCTGCAAAACAGA 717
      177  IleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuAlaAspLeuProAsn--- 195
      718  GATCATAGAGAGAGCAGCCCTGTCTCTCT---ProValIleProValAlaValIuTrpAsnGluMetSer 750
      196  -----CAGAGATGCCACCTCCAGCCAGCTGC----- 777
      751  -----AGCCCTCAGCAAAATATAGCCCTATAAATAC----- 810
      209  LeuAlaGlnGluLeuGlnSerProGlnAspCysIleGlnLeuLeuLysLeuIleArg 228
      778  -----AGCCCTCAGCAAAATATAGCCCTATAAATAC----- 810
      229  LeuProAsnIleProHisGlnCysTrpLeuThrLeuGlnTyrLeuLeuLysHisPhePhe 248
      810  ----- 810
      249  LysLeuSerGlnAlaSerSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePhe 268
      810  ----- 810
      269  SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGlnHisLeuIleLys 288
      811  -----ACAAGCTGG----- 819
      289  AlaIleGluIleLeuIleSerThrGluTrpAsnGluArgGlnProAlaProAlaLeuPro 308
      820  ---AGACCACCTTCCCC-----AAAAGTGTGATAGAAAGATGTC 858
      309  ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328
      859  CAGCACAATGATGTGATCTTGGAGAAATACAGCCCGCAGCAGTGAAGAAGCATTCATG 918
      329  GlnAspAlaGlnIuTrpIuTrpGlyAspIleSerArgIuValAsnMetLys-----Leu 346
      919  AAGGAGACAAGAGATGATGTTCTGTCGCGATGTTCCACAAATCCAAAGAGAG 978
      347  ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrIysMetHisGlyAsp 366
      979  CCTATGTTTGGCTGTGTTTATGAGAACAAAGTCTACAAATGTAA--ATCCGCTTC 1035
      367  ---TyrThrLeuThrLeu-----ArgIysGlyGlyAsnAsnLysLeuIleIysIle 382
      1036  CTGGAGAGAAATCAGCAGTTGGCCCTGGGACAGCAGCTCAGAGAGATGAGAACTTGAT 1095
      383  PheHisArgAspGlyLysIuTrpGlyPheSerAspProLeu-----ThrPheAsn 398

```

QY 1096 TCAGTAGAGACATCATGACACACTACAGAAAT 1128
||||| :||||| :||||| :||||| :
Db 399 SerValValGluLeuIleAsnHisTyrArgAsn 409

RESULT 9
US-08-391-615-2

Sequence 2, Application US/08391615
Patent No. 5550054

GENERAL INFORMATION:

APPLICANT: Witte, Owen

APPLICANT: Tsukada, Satoshi

APPLICANT: Saifran, Douglas

APPLICANT: Rawlings, David

TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE

TITLE OF INVENTION:

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/391,615

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/006,449

FILING DATE: 21-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Rowland, Berttram I

REGISTRATION NUMBER: 20,015

REFERENCE/DOCKET NUMBER: A-57583-1/BIR

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299 FHT UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 659 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-391-615-2

Alignment Scores:

Pred. No.: 2,996-06

Score: 143.00

Percent Similarity: 33.98%

Best Local Similarity: 23.44%

Query Match: 7.08%

Db: 1

Gaps: 11

US-09-856-061-3 (1-1129) x US-08-391-615-2 (1-659)

QY 373 AGAAGCCAAACATTAAGAGATGATCGTAAAGAAAGAAAGATTCCTTACCACT 432

||||| :||||| :||||| :||||| :
Db 171 ArgAsnGlySerLeuYsProGlySerSerHisArgLysThrLysLysProLeuPro 190

QY 433 CCTGGCCCTCTCATTAACACTCCGGAAGAAAGTACCAACCTGGCCCTTGACCGGAGAGC 492

||||| :||||| :||||| :||||| :
Db 191 ThrProGluGluAspGlnIleLeuLysLys-----ProLeuProProGluProHra 208

QY 493 AGAGGCCACCTTATCTCAGACACACACTTCCAGAGTCCAGGGAATGCCAGTCAG 552

||||| :||||| :||||| :||||| :
Db 209 -----AlaPro 210

QY 553 ATAACCTTAAGGACCTTAAGTGAGGCTCTTGAGAA---GAAAAGTTCCTCATTAACGAG 609
||||| :||||| :||||| :||||| :
Db 211 IleSerThrThrGluLeuLysLysValValAlaLeuYrAspTyrMetProMetAsnAla 230

QY 610 -----AGAAGCCCTGAATCACTCACTCTGTGTGAACCAACCAATATCACTCA 654
||||| :||||| :||||| :||||| :
Db 231 AsnAspLeuGlnLeuArgLysGlyGluLysPheIleLeuGluLysSerAsn----- 248

QY 655 GAGATTCACCTTGCCTTATAGCAGTCTTCATTCAGACACAGAACACACAGTGTGCAAAAC 714
||||| :||||| :||||| :||||| :
Db 249 ---LeuProTyr-----TyrPArgAla 254

QY 715 AGAGATCATTAAGAGAGACGATGAGCCCTGTCTCTCAGAGATGCCAGCTTCACCGAGC 774
||||| :||||| :||||| :||||| :
Db 255 ArgAspLysAsnGly----- 259

QY 775 TGCAGCCCTCAGCAAAATATACTGCCCTTAATATACACAGCTGACAGACACCTTCCCG 834
||||| :||||| :||||| :||||| :
Db 260 -----GlnGluGlyTyrIleProSerAsnTyrIleThr----- 270

QY 835 AAAAGTCTGATAGAAGAGATGTCGACACACATGAATGATGATGATGAGAAATACAGCCGC 894
||||| :||||| :||||| :||||| :
Db 271 -----GluAlaGluAspSerIleGluMetTyrGluIlePtyrSerLysHisMetThrArg 288

QY 895 CAGCAGTGTGAAGAGGACATTCATGAAGAGAAAGAGAGATGTTCTGTGTCGAGAT 954
||||| :||||| :||||| :||||| :
Db 289 SerGlnAlaGluLeuLeuLysGlnGluGlyGlyGluGlyGlyPheIleValArgAsp 308

QY 955 TGTTCACAAATTCACAGAACAGCCCTATGTGTTGGCTGTGTTTATGAGAAACAAAGTC 1014
||||| :||||| :||||| :||||| :
Db 309 SerSerLysAlaGlyLys----- 314

QY 1015 TACATGTAAATATCCGCTTCCTCGAGAGAGATACAGACATTTGCCCTGGAGACAGACTC 1074
||||| :||||| :||||| :||||| :
Db 315 TyrThrValSerVal-----PheAlaLysSerThrGly---- 325

QY 1075 AGAGAGATGAGAAAGTTGATTCAGTAGAAGACATCATGCAACACTAC 1122
||||| :||||| :||||| :||||| :
Db 326 -----GluProGlnGlyValIleArgHisTyr 334

RESULT 10
US-08-426-509A-8

Sequence 8, Application US/08426509A

Patent No. 6326469

GENERAL INFORMATION:

APPLICANT: Ulrich, Axel

APPLICANT: Gishizky, Mikhail

APPLICANT: Sures, Itman G.

TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN

TITLE OF INVENTION: TYROSINE KINASES

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/426,509A

FILING DATE: 21-APR-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/232,545

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

```

REFERENCE/DOCKET NUMBER: 7683-0074-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: No. 6326469e
US-08-426-509A-8

Alignment Scores:
Pred. No.: 6,12e-06
Score: 140.00
Percent Similarity: 33.98%
Best Local Similarity: 22.66%
Query Match: 6.93%
DB: 4
Caps: 11

US-09-856-061-3 (1-1129) x US-08-426-509A-8 (1-659)
QY 373 AGAAGCCAAACATTAAAGGATGATCCGTAAGAAAGACAGATTCTTACCACT 432
DB 171 ArgasnlyserleuylsproglyserSerHisArgylsThrlylsProleupPro 190
QY 433 CCTGGCCTCTCATACACTCCGAGAGATACCAACCTTGGCCCTGAGCCGGAAGC 492
DB 191 ThrProgluGluspglnlleleuylslys-----ProleupProgluProAla 208
QY 493 AGCAGGCCACCTTATCTCAGACACACCTTCCAGAGTCCAGGAATGCCAGTCAG 552
DB 209 -----AlaPro 210
QY 553 ATAAGCTTAAGGACTTAAGTGAAGTCTTGAAACA---GAAAAGTCTCTATACAG 609
DB 211 ValSerThrsergluLeuylsValAlaLeuylsPtyrMetPromelaAla 230
QY 610 -----AGGAAGCCTGAATCACTCATCTGTAGAAACCAAAATACCA 654
DB 231 AsnAspLeuInleuArglysglyAspGlyTyrPheIleleuInleuSerAsn----- 248
QY 655 GAATTCACACTTGCCATTACAGCTTCTTCATCAGACAGCAACCAAGCTGTCAAAAC 714
DB 249 ---LeuprOTrp-----TprArgAla 254
QY 715 AGAGATCATAGAGGAGCATGACCCCTGTTCTCAGAGATGCCAGCTCCAGCCAGC 774
DB 255 ArgAspLysasnly----- 259
QY 775 TGCAGCCCTCAGCAAAATATACTGCCCTTAAATACACAGCTGAGACACCTTCCCC 834
DB 260 -----GlnGluluglyTyrIleProserAsnTyrValThr----- 270
QY 835 AAAAGCTGATAGAAAGATGTCACACAAATGAATGTCATGAGAGATACAGCCGC 894
DB 271 -----GluAlaGluAspSerIleGluMetTyrGluTyrPtyrSerlyshMetThrArg 288
QY 895 CAGGCAAGGAGGATTCATGAGAGACAGCAAGATGTTCTTGTGCGCAGAT 954
DB 289 serGlnAlaGluInleuLeuylsGlnGluGlyLysGlnGlyGlyPheIleValAlaGasp 308
QY 955 TGTTCACCAAAATCCAGAGAGAGCCCTATGTTGGCTGTGTATAGAAACAAGTC 1014
DB 309 SerSerLysAlaGlyLys----- 314
QY 1015 TACAAATGTAATAATCCGCTTCTGAGAGAGATACAGACTTTGCCCTGGGACAGACTTC 1074
DB 315 TrrTrrHValSerVal-----PheAlaLysSerThrGly---- 325
QY 1075 AGAGGATGAGAGAGTTGATTGATTAGTAAGACATCATGACACACTAC 1122

```

```

DB 326 -----AspProGlnGlyValIleArgHisTyr 334

RESULT 11
PCT-US95-05008-8
Sequence 8, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-Apr-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-Apr-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-0074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-05008-8

Alignment Scores:
Pred. No.: 6,12e-06
Score: 140.00
Percent Similarity: 33.98%
Best Local Similarity: 22.66%
Query Match: 6.93%
DB: 5
Caps: 11

US-09-856-061-3 (1-1129) x PCT-US95-05008-8 (1-659)
QY 373 AGAAGCCAAACATTAAAGGATGATCCGTAAGAAAGACAGATTCTTACCACT 432
DB 171 ArgasnlyserleuylsproglyserSerHisArgylsThrlylsProleupPro 190
QY 433 CCTGGCCTCTCATACACTCCGAGAGATACCAACCTTGGCCCTGAGCCGGAAGC 492
DB 191 ThrProgluGluspglnlleleuylslys-----ProleupProgluProAla 208

```

```

OY 493 AGCAGGCCACCTTTATCTCAGAGACACACTTTCAGAAAGTCCAGGAAATGCCACGTGAG 552
DB 209 -----Alapro 210
OY 553 ATAGCTTAAGGAGCTTAAGTGGCTCCTTGAAGA--GAAAAAGTTCCTATACAGAG 609
DB 211 ValSerThrSerGluLeuLysGlyValAlaLeuTyrrAspTyrMetProMetAsnAla 230
OY 610 -----AGNAGCCTGAATCACTCATCTCTGTGTGAAGAACCAATACTCA 654
DB 231 AsnAspLeuGlnLeuArgLysGlyAspLysGlyPheLeuLeuGlnLysSerAsn----- 248
OY 655 GAGATTCACCTTGCATTAGCAGTCTTCTCATTCAGACAGCAACACACAGTGTGCAAAAC 714
DB 249 ---LeuProTrp-----TrpArgAla 254
OY 715 AGAGATCATAGAGAGGAGCATGCGACCTGTCTCTCCTCAGAGATGCCAGCTCCAGCCAGC 774
DB 255 ArgAspLysAsnGly----- 259
OY 775 TGCACCCCTCAGCAAAATATATCTGCTTAAATACACAGCTGGAGACACCTTTCCCC 834
DB 260 -----GlnGlnGlyTyrIleProSerAsnTyrValThr----- 270
OY 835 AAAAGCTGTATAGAAAGGATGTCACGACCAATGATGTATGATGGAGATACAGCCGC 894
DB 271 -----GluAlaGlnAspSerIleGluMetTyrGluTrpTyrSerLysHisMetThrArg 288
OY 895 CAGGACGTGAGAGAGGAGCATTCATGAAGAGAAAGAGATGTATGTTCTGTGTCGAGAT 954
DB 289 SerGlnAlaGlnGlnLeuLeuLysGlnGlnGlyGlyPheIleValArgAsp 308
OY 955 TGTTCACAAATCCAGAGAAAGCCCTATGTTTGGCTGTGTATATAGAAACAAAGTC 1014
DB 309 SerSerLysAlaGlyLys----- 314
OY 1015 TACAAATGTAAATCCGCTTCTCTCGAGAGAAATACAGCACTTGGCCCTGGAGACAGCATC 1074
DB 315 TyrThrValSerVal-----PheAlaLysSerThrGly----- 325
OY 1075 AGAGAGATGAGAAAGTTGATTCAGTAGAAGACATCATGCAACACTAC 1122
DB 326 -----AspProGlnGlyValIleArgHisTyr 334

RESULT 12
US-08-391-615-6
Sequence 6, Application US/08391615
GENERAL INFORMATION:
APPLICANT: Witte, Owen
APPLICANT: Tsukada, Satoshi
APPLICANT: Saffran, Douglas
APPLICANT: Rawlings, David
TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embardadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,615
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 08/006,449
FILING DATE: 21-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-57583-1/BJR UCLA 232-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-615-6

Alignment Scores:
Pred. No.: 1,03e-05 Length: 442
Score: 137.00 Matches: 60
Percent Similarity: 34.50% Conservative: 29
Best Local Similarity: 23.26% Mismatches: 67
Query Match: 6.78% Indels: 102
DB: 1 Gaps: 11

US-09-856-061-3 (1-1129) x US-08-391-615-6 (1-442)
OY 373 AGAAGCCAAAGATTAAAGAGATGATCGTATGAAGAAAGACAGATCTTATCCACCT 432
DB 171 ArgAsnGlySerLeuLysProGlnSerSerHisArgLysThrLysLysProLeuPro 190
OY 433 CCTGGCCTCTCATTAACACTTCCGAAGAAAGTACCAACCTTGCCCTGAGCCGGAGAGC 492
DB 191 ThrProGlnLysAspGlnLeuLysLys-----ProLeuProGlnProAlaAla 208
OY 493 AGCAGGCCACCTTATCTCAGAGACACACTTTCAGAGATC-----CAGGAGATG 543
DB 209 Ala-----ProValSerThrSerGluLeuLysLysGlnValAlaLeuTyrrAspTyrMet 226
OY 544 CCC-----AGTCAGATTAAGCTTAAGGACTTAAGTGAG-----GTCTTGAAGCA 588
DB 227 ProMetAsnAlaAsnAspLeuGlnLeuArgLysGlyAspLysTyrPheIleLeuGlnGlu 246
OY 589 GAAAAAGTCTCATTAACAGAGAGAGCCGTGATCAATCATCTGTGTAAGAAACCAAAAT 648
DB 247 SerAsnLeuProTrpTrpArgAlaArg-----AspLysAsn 258
OY 649 ACTCAGAGATTCACCTTGCATTAGCAGTCTTCAATTCAGACAGCAACACCACTGTG 708
DB 259 GlyGlnGlu-----GlyTyrIleProSerAsnAspValThr 270
OY 709 CAAACAGAGATCATAGAGAGAGAGCAGCCCTGTCTCCTCAGAGATGCCAGCTCCA 768
DB 271 GluAlaGlnAsp----- 274
OY 769 GCCAGCTGACGCCCTCAGCAAAATATATAGTCCCTATTAATACAAAGCTGGAGACCACT 828
DB 274 ----- 274
OY 829 TTCCCCAAAGGTCTGATAGAAAGATGCCAGACATGATGTATGAGAAATAC 888
DB 275 -----SerIleGluMetTyrGluTrpTyrSerLysHisMet 286
OY 889 AGCCGCGAGCAGCTGAGAGAGGAGCATTCATGAAGAGAAAGAGATGATGTTCTGTGTC 948
DB 287 ThrArgSerGlnAlaGlnGlnLeuLeuLysGlnGlnGlyGlyPheIleVal 306
OY 949 CGAGATTTGTTCCACAAATCCAGAGAAAGAGCCCTATGTTTGGCTGTGTTTATGAGAAC 1008
DB 307 ArgAspSerSerLysAlaAlaLys-----TyrThrLeuSerVal----- 319

```

OY 1009 AAGCTCAGCAATGTAAATCCGCTTCTCGAGAGGAATCAGCACTTGCCTGGGACA 1068
DB 320 -----PheAlalysSerThr 324
OY 1069 GACCTACAGAGGATGAGACGTTGATTCAGTAGAAGCATCAGCACTAC 1122
DB 325 Gly-----AspProGlnGlyValIleArgHisTyr 334
RESULT 13
US-08-202-389-28
Sequence 28, Application US/08202389
Patent No. 5536636
GENERAL INFORMATION:
APPLICANT: Freeman Jr., Robert M.
APPLICANT: Plutzky, Jorge
APPLICANT: Neel, Benjamin G.
APPLICANT: Rosenberg, Robert D.
TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,389
FILING DATE: 28-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,926
FILING DATE: 01-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/829,141
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/721,112
FILING DATE: 26-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: BI92-05MA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-202-389-28
Alignment Scores:
Pred. No.: 2.52e-05 Length: 92
Score: 130.00 Matches: 26
Percent Similarity: 64.71% Conservative: 26
Best Local Similarity: 34.12% Mismatches: 22
Query Match: 6.44% Indels: 8
DB: 1 Gaps: 3
US-09-856-061-3 (1-1129) x US-08-202-389-28 (1-92)
OY 871 TGGTACATTTGGGAATACGCCCGCAGCGCATGTGAAGAGCATTCATGAAGAGAACAAG 930
DB 1 TrrpHnIslgLyAsnLeuSerGlyLysGluAlaGluLysLeuIleLeuLunArgLys 20

OY 991 GATGTAGTTCTTGCTCCGAGATTTGTCACAAATCCAGAAGACCCTATGTTTG 990
DB 21 AsnGlySerPheLeuValArgGlu-----SerGlnSerLysProGlyAspPheValLeu 38
OY 991 GCTGTGTTTATGAGAACCAAGCTCTACATGTAAATCCGCTTCCTCGAGAGCAATCG 1050
DB 39 SerValArgThrAspAspLysValThrHisValMetIleArgTyr-----GlnAspLys 56
OY 1051 CAGTTTCCCTGGGACAGCACTCAGAGAGATGAGAGTTGATTCAGTAGAAGACATC 1110
DB 57 LysTyrAspValGly-----GlyGlyGlnSerPheGlyThrLeuSerGluLeu 72
OY 1111 ATCGAACCTACGAAG 1125
DB 73 IleAspHisTyrLys 77
RESULT 14
US-08-729-416C-1
Sequence 1, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-1
Alignment Scores:
Pred. No.: 6.36e-05 Length: 474
Score: 129.50 Matches: 72
Percent Similarity: 37.33% Conservative: 37
Best Local Similarity: 24.66% Mismatches: 112
Query Match: 6.41% Indels: 71
DB: 3 Gaps: 16
US-09-856-061-3 (1-1129) x US-08-729-416C-1 (1-474)
OY 334 AGTTGGAAGAGATGAGCAACCCATTTCCAGAGAGTCAGAAAGCCAAATTAAGA 393
DB 201 ArgMetGlnSerLeuAspLysProTrrpThrGluGlu-----GluGly 214
OY 394 GATGCATCCGTAAAGAAAG-----AACAAAGATTCTTACACACCTCCGCGCCTGCATATA 447

```

Db 215 AspGlySerAspHisProTyrTyrAsnSerLeuProSerLysMetProProGlyGly 234
448 AACCTCCGCAAGAAAGTACCAACCCCTGGCCCGGAGAGCAGC-----495
Db 235 PheLeuAspThrLysLeuLysProArgProHisAlaProAspThrLysLysLysLys 254
496 -----AGGCCACCTTTATCTCAGAGAC-----ACCTTCCAGAA---GTCCAG 537
Db 255 LysGlnGlnThrTyrTyrGlnGlnArgHisLeuGlyAspThrPheGlyLysAspTrpGln 274
538 GGAAATCCCAAGTCAGATTAAGCTTAAGGACTTAAGTGGAGCTCTTGAAGCAGAA---AAA 594
Db 275 GlnThrProLeuArgLysLysSerAspLysSerThrProGlnGlyLysLeuHis 294
595 GTTCCCATTAACAGAGAGAGCCTGAATCAGCATCTGTATAAAACCAAAATACTCA 654
Db 295 ValAlaProThrGlyLysAlaPro-----ThrTyrVal-----AsnThrGln 308
655 GAGATT-----CCACTTGCATTAAGCATTCTTCATTACAGAGAGCAGAAC 699
Db 309 GlnIleProGlnIleAlaTrpProAlaAlaValSerSerLysLysLysLysLys 328
700 CACAGTGTGCAGAAACAGAGATCATAGAGAGCAGT-----CAGCCCTGTCTCCT-- 750
Db 329 AspLeuPheAspMetLysProPheGlnAspAlaLeuLysAsnGlnProLeuGlyProVal 348
751 -----CAGAGATGCCAGCTCCAGCCAGCTGCGAGCCCTCAGAGAAAT 792
Db 349 LeuSerLysAlaAlaSerValGlyLysLysSerProValSer-----362
793 ATACTGCCCTATAATATACACAGCTGGAGCACCTTCCCAAAAGCTGTATAGAAAG 852
Db 363 -----ProArgAlaProAspAlaLysMetLeuGln 372
853 GATGTCACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 912
Db 373 GlnLeuGlnIleLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 392
913 TTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
Db 393 Leu-----GlnLysAspGlyAspPheLeuValArg-----LysSerThr 405
973 GAGAGAGC-----TATGTTGGCTGCTGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1026
Db 406 ThrAsnProGlySerPheValIleThrGlyMetHisAsnGlyLysLysLysLysLys 425
1027 ATCCGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086
Db 426 LeuValAspProGln-----GlyThrLysLysLysLysLysLysLysLysLysLys 438
QY 1087 AAGTTGATTCAGTACAGACATCATCGAACAATAC 1122
Db 439 ValPheAspSerLysSerHisLeuLeuAsnHisHis 450

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-7
Alignment Scores:
Pred. No.: 7,11e-05 Length: 594
Score: 129.50 Matches: 72
Percent Similarity: 37.33% Conservative: 112
Best Local Similarity: 24.66% Mismatches: 71
Query Match: 6.41% Gaps: 16
DB: 3
US-09-856-061-3 (1-1129) x US-08-729-416C-7 (1-594)
QY 334 AGGTGGAAGAGTGCAGAAACCATTTCCAGGAGCCTGAGAAACCAATTAAGCA 393
Db 321 ArgMetIleSerLeuAspGlnProTrpThrGlnGln-----GlnGly 334
QY 394 GATGTCACAGCAGTAAAGAAAG-----AACAGATTCCTTACACCTCCGCGCTCATA 447
Db 335 AspGlySerAspHisProTyrTyrAsnSerLeuProSerLysMetProProGlyGly 354
448 AACCTCCGCAAGAGTACCAACCCCTGGCCCGGAGAGCAGC-----495
Db 355 PheLeuAspThrLysLeuLysProArgProHisAlaProAspThrLysLysLysLys 374
496 -----AGGCCACCTTTATCTCAGAGAC-----ACCTTCCAGAA---GTCCAG 537
Db 375 LysGlnGlnThrTyrTyrGlnGlnArgHisLeuGlyAspThrPheGlyLysAspTrpGln 394
538 GGAAATCCCAAGTCAGATTAAGGACTTAAGTGGAGCTCTTGAAGCAGAA---AAA 594
QY 395 GlnThrProLeuArgLysLysSerAspLysSerThrProGlnGlyLysLeuHis 414
595 GTTCCCATTAACAGAGAGAGCCTGAATCACTCATCTGTATAAAACCAAAATACTCA 654
Db 415 ValAlaProThrGlyLysAlaPro-----ThrTyrVal-----AsnThrGln 428
655 GAGATT-----CCACTTGCATTAAGCATTCTTCATTACAGAGAGCAGAAC 699
Db 429 GlnIleProGlnIleAlaTrpProAlaAlaValSerSerLysLysLysLysLys 448
700 CACAGTGTGCAGAAACAGAGATCATAGAGAGCAGT-----CAGCCCTGTCTCCT-- 750
Db 449 AspLeuPheAspMetLysProPheGlnAspAlaLeuLysAsnGlnProLeuGlyProVal 468
751 -----CAGAGATGCCAGCTCCAGCCAGCTGCGAGCCCTCAGAGAAAT 792
Db 469 LeuSerLysAlaAlaSerValGlyLysLysSerProValSer-----482
793 ATACTGCCCTATAATATACACAGCTGGAGAGCAGCTTCCCAAAAGCTGTATAGAAAG 852
Db 483 -----ProArgAlaProAspAlaLysMetLeuGln 492

```

```
OY 853 GATGTCACACACATGATGTACATTGGAGATACAGCCGACGACAGAGAGCA 912
      :::::||||| ||||| ||||| ||||| |||
Db 493 GluLeuGlnAlaGluThrTrpTyrGlnGlyGluMetSerArgIbGlnAlaGluGlyLeu 512
OY 913 TTCATGAGAGAGACACAGATGCTACTTCTTGTCGAGATTGTTCCACAAAATCCAG 972
      ||||| ||||| ||||| ||||| |||||
Db 513 Leu-----GluLysAspGlyAspPheLeuValArg-----LysSerThr 525
OY 973 GAAGAGCCC-----TATGTTTGGCTGCTTTATGACACAAAGCTACATGTAAAA 1026
      ||| ::||| ||| ::|
Db 526 ThrAsnProGlySerPheValLeuThrGlyMetHisAsnGlyGlnAlaLysHisLeu 545
OY 1027 ATCCGCTTCCTGGAGAGGAATCAGAGTTGCCCTGGGGACAGACTCAGAGAGATGAG 1086
      ::| ||| ::| ||| ||| ||| |||
Db 546 LeuValAspProGlu-----GlyThrIleArgThrLysAspArg 558
OY 1087 AAGTTGATTCAGTAGACACATCATCCAGACCTAC 1122
      ||||| ||||| ||||| ||||| |||||
Db 559 ValPheAspSerIleSerHisLeuIleAsnHisHis 570
```

Search completed: April 21, 2003, 12:33:09
Job time : 25.2707 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucletc - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:20:58 ; Search time 28.7202 Seconds

(without alignments)
5944.816 Million cell updates/sec

Title: US-09-856-061-3

Perfect score: 2020
Sequence: 1 ttccagaacttcagctgcgc.....catcgacactacagaatt 1129

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 288829 seqs, 7561385 residues

Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO/US0985661/runat.21042003_113021_3809/app.query.fasta.1.3150
-DB=Published.Applications_AA -OFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdt -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=local -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0985661.ecgn.1.1.34_etrunat.21042003_113021_3809
-NCPU=6 -ICPU=3 -NO_XLPYX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMECUT=120 -WARN_TIMECUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PTC_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PTCUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2009	99.5	428	9	US-09-966-955A-4
2	2009	99.5	443	9	US-09-966-955A-2
3	1694	83.9	353	9	US-09-966-955A-6
4	238	11.8	46	10	US-09-864-761-38652

5	214	10.6	40	10	US-09-864-761-44023	Sequence 44023, A
6	214	10.6	40	10	US-09-864-761-44024	Sequence 44024, A
7	166	8.2	30	10	US-09-864-761-38653	Sequence 38653, A
8	150.5	7.5	724	10	US-09-962-925-4	Sequence 4, Appl1
9	149	7.4	688	9	US-10-081-980B-1	Sequence 1, Appl1
10	149	7.4	724	9	US-10-081-980B-3	Sequence 1, Appl1
11	147	7.3	724	10	US-09-962-929-2	Sequence 2, Appl1
12	144.5	7.2	945	8	US-08-965-927-2	Sequence 2, Appl1
13	144.5	7.2	945	12	US-10-121-882-2	Sequence 2, Appl1
14	143	7.1	659	9	US-10-045-202-4	Sequence 4, Appl1
15	140	6.9	659	9	US-09-977-260-8	Sequence 8, Appl1
16	140	6.9	659	9	US-10-045-202-2	Sequence 2, Appl1
17	140	6.9	659	9	US-09-977-261-8	Sequence 8, Appl1
18	140	6.9	659	10	US-09-977-269-8	Sequence 8, Appl1
19	126	6.2	597	9	US-10-038-010-22	Sequence 22, Appl1
20	125.5	6.2	533	10	US-09-908-805B-63	Sequence 63, Appl1
21	125	6.2	533	10	US-09-920-021A-3	Sequence 3, Appl1
22	120	5.9	939	10	US-09-926-248B-29	Sequence 29, Appl1
23	120	5.9	939	10	US-09-801-368-380	Sequence 380, App
24	119	5.9	620	9	US-09-977-260-9	Sequence 9, Appl1
25	119	5.9	620	9	US-09-977-261-9	Sequence 9, Appl1
26	119	5.9	620	10	US-09-977-269-9	Sequence 9, Appl1
27	115	5.7	675	9	US-10-186-399-3	Sequence 3, Appl1
28	115	5.7	675	9	US-09-977-260-4	Sequence 4, Appl1
29	115	5.7	675	9	US-09-977-261-4	Sequence 4, Appl1
30	115	5.7	675	10	US-09-977-269-4	Sequence 4, Appl1
31	112.5	5.6	552	10	US-09-880-192-50	Sequence 50, Appl1
32	112	5.5	1325	10	US-09-864-761-35612	Sequence 35612, A
33	111	5.5	482	10	US-09-764-853-568	Sequence 568, App
34	111	5.5	542	10	US-09-908-805B-44	Sequence 44, Appl1
35	111	5.5	655	10	US-09-205-658-57	Sequence 57, Appl1
36	111	5.5	655	10	US-09-844-353A-57	Sequence 57, Appl1
37	111	5.5	944	10	US-09-964-238-2	Sequence 2, Appl1
38	110	5.4	562	9	US-09-925-299-869	Sequence 869, App
39	110	5.4	562	10	US-09-925-299-869	Sequence 869, App
40	110	5.4	1341	9	US-10-076-622-565	Sequence 565, App
41	110	5.4	1341	12	US-10-076-622-565	Sequence 565, App
42	110	5.4	1349	9	US-10-076-622-573	Sequence 573, App
43	110	5.4	1349	12	US-10-076-622-573	Sequence 573, App
44	109.5	5.4	4019	9	US-09-854-133-425	Sequence 425, App
45	109.5	5.4	4019	10	US-09-738-973-425	Sequence 425, App

ALIGNMENTS

RESULT 1
US-09-966-955A-4
Sequence 4, Application US/09966955A
Patent No. US20020155563A1
GENERAL INFORMATION:
APPLICANT: Perez-Villar, Juan J.
APPLICANT: Chang, Han
APPLICANT: Yang, Wen-Pin
APPLICANT: Wu, Yuli
APPLICANT: Whitney, Steven S.
TITLE OF INVENTION: Identification and Cloning of a Full-Length Human
TITLE OF INVENTION: Clnk-related Gene, MIST (Mast Cell Immoreceptor
TITLE OF INVENTION: Signal Transducer)
FILE REFERENCE: 3053-4113051
CURRENT APPLICATION NUMBER: US/09/966, 955A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/237030
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 428
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #7, AMINO ACID
OTHER INFORMATION: SEQUENCE


```

Db 114 ThrietyrPhelysValAlaMetAspThrProLeuProLeuAspThrArgThrSerIle 133
QY 301 TCCATGGAGACGCCGACCTGGAGACAGAGAGTGGGAAGAGTGGACCAACCATTT 360
Db 134 SerIleGlyLeuProThrTrpAsnThrGlnThrArgLeuGlnArgValAspIle 153
QY 361 TCCAGGAGCTCAGAGCCCAAAACATTAAAGAGATGCATCCCTAGAGAAAGAACAGATT 420
Db 154 SerIleAspValArgSerGlnAsnIleIysGlyAspAlaSerValArgIleAsnIle 173
QY 421 CCTTACACCTCTCCGCTCTCATACACTTCGAGAGAGTACCAACCTTCCCTCCCT 480
Db 174 ProLeuProProArgProLeuIleThrLeuProIleIysIleGlnProIleProPro 193
QY 481 GACCCGAGAGCAGAGCGACCTTATCTACAGACACACCTTCCAGAGTCCAGGGA 540
Db 194 GluProGlnSerSerArgProProLeuSerGlnArgHisThrPheProGluValGlnArg 213
QY 541 ATGCCAGTCAGATAAGCTTAAGGAGACTTAAGTAGAGTCCCTGAGAGCAAAAGTCTCT 600
Db 214 MetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnIleValPro 233
QY 601 CATACCCAGAGAGACCTGATCACTCATCTGTAGAAAACCAAAATACTCAAGAGATT 660
Db 234 HisAsnGlnArgIysProGlnSerThrHisLeuGlnAsnGlnAsnThrGlnGluIle 253
QY 661 CCACCTGGCCATTAGCAGTCTTCATTCAGAGCAAGACAGCAGTGGCAAAACAGAGAT 720
Db 254 ProLeuAlaIleSerSerSerSerPheThrThrSerAsnHisSerValGlnAsnArgAsp 273
QY 721 CATAGAGAGAGCATGACGCCCTGTTCTCTCAGAGATGCCAGCTCCAGCAGCTGCAGC 780
Db 274 HisArgGlyGlyMetGlnProCysSerProGlnArgCysGlnProAlaSerCysSer 293
QY 781 CCGCAGCAAAATATACGCCCCCTAATAATACACAAGCTGGAGACCACTTCCCAAAAG 840
Db 294 ProHisGlnAsnIleLeuProIleIysIleIysIleIysIleIysIleIysIleIysIle 313
QY 841 TCTGATGAAGAGATGTCACGACCAATGAATGTGATTCAGAGAAATACACGCCGACGCA 900
Db 314 SerAspArgIysAspValGlnHisAsnGlnIleIleIleIleIleIleIleIleIleIle 333
QY 901 GTGAGAGAGCATTCATGAAGAGAGACAGAGATGCTTCTTGTCCGAGATTGTTC 960
Db 334 ValGlnIleAlaPheMetIysGlnAsnIysAspIleSerPheLeuValArgAspCysSer 353
QY 961 ACAAAATCCAGAGAGAGCCCTATGTTGGCTGTGTTATGAGACCAAGTCTACAT 1020
Db 354 ThrIysSerIysGlnGlnProIleIleIleIleIleIleIleIleIleIleIleIleIle 373
QY 1021 GTAAATAATCCGCTCTCGAGAGAGAAATCAGAGATTGCCCTGGGAGAGAGACTCAGAGA 1080
Db 374 ValIysIleArgPheLeuGlnArgAsnGlnGlnPheIleAlaLeuGlnIleGlnIleArgIle 393
QY 1081 GATGAGAAGTTGATTGATGAGTAAGACATCATCGACACTCAACAAGAT 1128
Db 394 AspIleIysPheAspSerValGlnAspIleIleGlnIleIleIleIleIleIleIleIleIle 409

```

RESULT 3

```

US-09-966-955A-6
; Sequence 6, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitley, Gena S.
; APPLICANT: Kanter, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-Length Human
; TITLE OF INVENTION: Clink-related Gene, MIST (Mast Cell Immunoceptor)
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1

```

```

; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 6
; LENGTH: 353
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPLICE VARIANT CLONE #12, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-6

```

Alignment Scores:

```

Pred. No.: 9,24e-136 Length: 353
Score: 1694.00 Matches: 317
Percent Similarity: 99.69% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 83.86% Indels: 0
DB: Gaps: 0

```

US-09-856-061-3 (1-1129) x US-09-966-955A-6 (1-353)

```

QY 172 ATGAGAGAGACATGCGAGTGAATTAATTTACAGCCCGGCTATTAAGAAATCTGAA 231
Db 1 MetGlnGlnThrTrpGlnSerIleIysIleLeuProAlaArgProIleIysGlnSerGlu 20
QY 232 TATGACATACACACTATTTCAGAGTTGCAATGACACTCCCTCCGTTAGACACCAGG 291
Db 21 TyrAlaAspThrHisIleIysIleValAlaMetAspThrProLeuProLeuAspThrArg 40
QY 292 ACCGTATCTCCATTGGACAGCCGACCTGAGACACACAGAGAGTGAAGAGTGCAG 351
Db 41 ThrSerIleSerIleGlyGlnProThrTrpAsnThrGlnThrArgLeuGlnArgValAsp 60
QY 352 AAACCCATTTCCAGGAGCTCAGAAAGCAAAACATTAAAGAGATGCATCCGTAGAAAG 411
Db 61 LysProIleSerIysAspValArgSerGlnAsnIleIysGlyAspAlaSerValArgIys 80
QY 412 AACAGATTCCTTACACACTCCCGGCTCTCATTAACCTTCGAGAAAGTACCAACC 471
Db 81 AsnIysIleProLeuProProProArgProLeuIleThrLeuProIysIleIysIleIysIle 100
QY 472 TTGCCCCCTGAGCGAGAGACAGAGCCACTTATCTCAGAGACACACTTCCAGAA 531
Db 101 LeuProProGlnProGlnSerSerArgProProLeuSerGlnArgHisThrPheProGlu 120
QY 532 GTCCAGGGAATCCCACTGATGATTAAGGAGCTTAAGGAGTTCCTTGAAGCAGAA 591
Db 121 ValGlnArgMetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnValGlu 140
QY 592 AAATCTCTCATACACAGAGAGACCCGATCACTCATCTGTTAGAAACCAAAATACT 651
Db 141 LysValProHisAsnGlnAlaGlyAspProGlnSerThrHisLeuGlnAsnIleIleIle 160
QY 652 CAAGAGATTCACCTTGCATTAAGAGTCTTCATTCAGAGCAAGCAACACAGAGTGCAG 711
Db 161 GlnIleIleProLeuAlaIleSerSerSerSerPheThrThrSerAsnHisSerValGln 180
QY 712 AACAGATCTATAGAGAGAGCATCAGCCCTGTTCTCTCAGAGATGCCAGCTCCAGCC 771
Db 181 AsnArgAspHisArgIysGlyIleMetGlnProCysSerProGlnArgCysGlnProProAla 200
QY 772 AGCTGAGCCCTCAGAGAAATATACGCTGATTAATTAACACCACTGAGAGCACTTTC 831
Db 201 SerCysSerProHisGlnAsnIleLeuProIleIysIleIysIleIysIleIysIleIysIle 220
QY 832 CCCAAAGGTTGTATGAAGAGATGTCAGAGCAATGATGATGATGATGATGATGATGATGAT 891
Db 221 ProIysArgSerAspArgIysAspValGlnHisAsnGlnIleIleIleIleIleIleIleIle 240

```

```

; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

1	PRIOR	FILLING DATE:	2001-01-30	PCT/US01/00665
2	PRIOR	APPLICATION NUMBER:	PCT/US01/00665	
3	PRIOR	FILLING DATE:	2001-01-30	PCT/US01/00668
4	PRIOR	APPLICATION NUMBER:	PCT/US01/00668	
5	PRIOR	FILLING DATE:	2001-01-30	PCT/US01/00663
6	PRIOR	APPLICATION NUMBER:	PCT/US01/00663	
7	PRIOR	FILLING DATE:	2001-01-30	PCT/US01/00662
8	PRIOR	APPLICATION NUMBER:	PCT/US01/00662	


```

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38653
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005599.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
US-09-864-761-38653

Alignment Scores:
Pred. No.: 3,22e-06 Length: 30
Score: 166.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8,22% Indels: 0
DB: 10 Gaps: 0

US-09-856-061-3 (1-1129) x US-09-864-761-38653 (1-30)
QY 932 ATGTGAGTTTCTGTGTCGACATGTTCCACAAATCCAGAGAGCCCATGTTTGG 991
DB 1 MetValValSerTPSerGluIleValIProGlnAsnProArGlySerProMetPheTrp 20
QY 992 CTGTGTTTATGAGAAAGCTAGACATG 1021
DB 21 LeucysPheMetArGlyThrLysSerThrMet 30

RESULT 8
US-09-962-929-4
; Sequence 4, Application US/09962929
```

```

; Patent No. US20020115058A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; APPLICANT: Nielsen, Anne A.
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Diseases Associated wit
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/962,929
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-962-929-4

Alignment Scores:
Pred. No.: 0.000147 Length: 724
Score: 150.50 Matches: 82
Percent Similarity: 37.04% Conservative: 48
Best Local Similarity: 23.36% Mismatches: 93
Query Match: 7,45% Indels: 128
DB: 10 Gaps: 17

US-09-856-061-3 (1-1129) x US-09-962-929-4 (1-724)
QY 406 AGAAGACAAAGATCTCTTACCACTCCGCGCTTCATACACTTCGGAAGATAC 465
DB 80 LysylsIleSerProPheThrProLysProArPro-----ProArG----- 93
QY 466 CAACCTTGCCCTGCGCCGAGACGACGAG----- 498
DB 94 --ProLeuProValAlaProGlySerSerLysThrGluAlaAspValGluGlnAla 112
QY 499 -----CCACCTTATCTCAGAGACACACTTTCAGAGATC----- 534
DB 113 LeuThrLeuProAspLeuAlaGluGlnPheAlaProAspIleAlaProLeuLeu 132
QY 535 -----CAGGAATG----- 543
DB 133 IleLysLeuValGluAlaIleGluLysLysGlyLeuGlyCysSerThrLeuTyrArgThr 152
QY 544 CCCAGTCGATTAACCTTAAAGGACTTAAGTACGTCCTGAGACAGAAAGTTCCAT 603
DB 153 GlnSerSerSerAsnLeuAlaGluLeuArgGlnLeuLeuAspCysAsp---ThrProSer 171
QY 604 AACGAGAGAGACCTGATCAACTCATCTGTTAGAAACCA----- 645
DB 172 ValAspLeuGluMetIleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuLeu 191
QY 646 -----AATACTCAAGAGATTCACCTTGCCATTAGCAGTTCT----- 681
DB 192 AspLeuProAsnProValIleProAlaAlaValTyrSerGluMetIleSerLeuAlaPro 211
QY 681 ----- 681
DB 212 GluValGlnSerSerGluGluTyrIleGlnLeuLeuLysLeuIleArgSerProSer 231
QY 682 -----TCA 684
DB 232 IleProHisGlnTyrTrpLeuThrLeuGlnTyrIleLeuLysHisPhePheLysLeuSer 251
QY 685 TTCAGCAAGCAACCAAGTGTGCAAAACAGAGATCATAGAGAGGATGAGCCCTGT 744
DB 252 GlnThrSerSerLysAsnLeuLeuAsnAlaArg-ValLeuSerClnIlePheSerProme 271
QY 745 TCTCCTCAGA-----GATCCAGGCTCCAGCCAGCTGCAGCCCTCAGCAAA---ATA 795
DB 271 tLeuPheArgPheSerAlaAlaSerSerAspAsnThrGluAsnLeuIleLysValIleGln 291
```


/ CURRENT FILING DATE: 2002-08-23
 / PRIOR APPLICATION NUMBER: 60/270,014
 / PRIOR FILING DATE: 2001-02-20
 / NUMBER OF SEQ ID NOS: 12
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO: 3
 / LENGTH: 724
 / TYPE: PRT
 / ORGANISM: mus musculus
 / US-10-081-980B-3

Alignment Scores:
 Pred. No.: 0.000197 Length: 724
 Score: 149.00 Matches: 76
 Percent Similarity: 32.76% Conservative: 39
 Best Local Similarity: 21.65% Mismatches: 96
 Query Match: 7.38% Indels: 140
 DB: Gaps: 17

US-09-856-061-3 (1-1129) x US-10-081-980B-3 (1-724)

QY 421 CCTTACACACCTGCGGCTCTCATACACTTCGAGAGAGTACCAACCTTGCCCT 480
 DB 85 ProthProlysProAlaArgPro-----ProArg-----ProLeuProAla 97
 QY 481 GAGCGGAGAGCAGCAGC-----CCACCT 504
 DB 98 AlaProGlySerSerLysThrGluAlaAspThrGluGlnGlnAlaLeuProLeuProAsp 117
 QY 505 TTATCTCAGAGACACACCTTTCAGAGATCCAGGAGATGCCAGTACAGATAAGCTTAAG 564
 DB 118 LeuAlaGlnGlnInPheAlaProProAspVal--AlaProProLeuLeuIleLysLeuLeu 136
 QY 555 GACTTAAGTGAAGTCTCTTGAAGCAGAAAAAGTTCCTCATACACAGAGAGAGCT 618
 DB 137 GluAlaIleGlnLysLysGlnLysLysSerThrLeuTyrArgThrGlnSerSer 156
 QY 619 GAATCAACTCATCTGTAGAA-----AACCAAAATACTCAAGAG 657
 DB 157 AsnProAlaGlnLeuAlaArgGlnLeuLeuAspCysAspAlaAlaSerValAspLeuGlnMet 176
 QY 658 ATTCCACTTGCATTAAGAGATCTCTCATTCACAGACACACACAGTGTCCAAACAGCA 717
 DB 177 IleAspValIleValLeuAlaAspAlaPheLysArgTyrLeuAlaAspLeuProAsn-- 195
 QY 718 GATCATAGAGAGAGCAGCAGCTGTCTCT-----ProValIleProValAlaValTyrAsnGlnMetSer 750
 DB 196 -----ProValIleProValAlaValTyrAsnGlnMetSer 750
 QY 751 -----CAGAGATGCCAGCTCCAGCCAGCTGC----- 777
 DB 209 LeuAlaGlnGlnLeuGlnSerProGlnAspCysIleGlnLeuLysLysLeuIleArg 228
 QY 778 -----ACCCCTCAGAGAAATATACTGCCCTATANAATAC----- 810
 DB 229 LeuProAlaIleProIleGlnCysLysTrpLeuThrLeuGlnTyrLeuLysHisPhePhe 248
 QY 810 ----- 810
 DB 249 LysLeuSerGlnAlaSerSerLysAsnLeuLeuAsnAlaArgValLeuSerGlnIlePhe 268
 QY 810 ----- 810
 DB 269 SerProValIleuPheArgPheProAlaAlaSerSerAspAsnThrGlnHisLeuIleLys 288
 QY 811 -----ACAAGCTG----- 819
 DB 289 AlaIleGlnIleLeuIleSerThrGlnTyrAsnGlnLysGlnProAlaProAlaLeuPro 308
 QY 820 ---AGACCACCTTCCCC-----AAAAGTCTGATAGAAAGATGTC 858
 DB 309 ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328

QY 859 CAGCAACATGAATGCTACATTGGAGAAATACACCCCGCAGGAGCTGAAGAGCATTCATG 918
 DB 339 GlnAspAlaGlnLysTrpLysTrpLysIleSerArgGlnValAsnGlnLys--Leu 347
 QY 919 AAGGAGAACAGGATGATGTTCTTGTCGAGATGTTCCACAAATCCAGAGAAAG 978
 DB 348 ArgAspThrAlaAspGlnTyrPheLeuValAlaGspAlaSerThrLysMetHisGlnLys 367
 QY 979 CCTATGTTTGGCTGTGTTTATGAGAACAAAGTCTGACAAATGTAAAA---ATCCGCTTC 1035
 DB 368 ---TyrThrLeuThrLeu-----ArgLysGlnGlnLysAsnLysLeuIleLysIle 383
 QY 1036 CTGAGAGGAATACACACTTTCCTGGGAGACAGACTCAGAGAGATGAGAAATTGAT 1095
 DB 384 PheHisArgAspGlnLysTyrGlnLysSerThrLeuProLeu-----ThrPheAsn 399
 QY 1096 TCAGTAGAAGACATCATCAGACATCAAGAT 1128
 DB 400 SerValValGlnLeuLeuAsnHisTyrArgAsn 410

RESULT 11

/ Sequence 2, Application US/09962929
 / Patent No. US20020115058A1
 / GENERAL INFORMATION:

/ APPLICANT: Pedersen, Finn S.
 / APPLICANT: Soerensen, Annette B.
 / TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated wit
 / FILE REFERENCE: A-70004/RMS/DCF
 / CURRENT APPLICATION NUMBER: US/09/962,929
 / PRIOR APPLICATION NUMBER: US 09/666,644
 / PRIOR FILING DATE: 2000-09-22
 / NUMBER OF SEQ ID NOS: 4
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO: 2
 / LENGTH: 724
 / TYPE: PRT
 / ORGANISM: Mus musculus
 / US-09-962-929-2

Alignment Scores:
 Pred. No.: 0.000292 Length: 724
 Score: 147.00 Matches: 76
 Percent Similarity: 32.48% Conservative: 38
 Best Local Similarity: 21.65% Mismatches: 97
 Query Match: 7.28% Indels: 140
 DB: Gaps: 17

US-09-856-061-3 (1-1129) x US-09-962-929-2 (1-724)

QY 421 CCTTACACACCTGCGGCTCTCATACACTTCGAGAGAGTACCAACCTTGCCCT 480
 DB 85 ProthProlysProAlaArgPro-----ProArg-----ProLeuProAla 97
 QY 481 GAGCGGAGAGCAGCAGC-----CCACCT 504
 DB 98 AlaProGlySerSerLysThrGluAlaAspThrGluGlnGlnAlaLeuProLeuProAsp 117
 QY 505 TTATCTCAGAGACACACCTTTCAGAGATCCAGGAGATGCCAGTACAGATAAGCTTAAG 564
 DB 118 LeuAlaGlnGlnInPheAlaProProAspVal--AlaProProLeuLeuIleLysLeuLeu 136
 QY 565 GACTTAAGTGAAGTCTCTTGAAGCAGAAAAAGTTCCTCATACACAGAGAGAGCT 618
 DB 137 GluAlaIleGlnLysLysGlnLysLysSerThrLeuTyrArgThrGlnSerSer 156
 QY 619 GAATCAACTCATCTGTAGAA-----AACCAAAATACTCAAGAG 657
 DB 157 AsnProAlaGlnLeuAlaArgGlnLeuLeuAspCysAspAlaAlaSerValAspLeuGlnMet 176

Query Match:

```

Length: 659
Matches: 60
Conservative: 27
Mismatch: 71
Indels: 98

```



```

Db 140 GLUAsp-----AspAlaAspTyrGluProProSerAsnAspGlu 153
OY 178 GAGACATGGCAGTGGATTAATAATTTTACACCGCGCTATAAAGATCGATATGCA 237
Db 154 GluAlaLeuGlnAsn---SerIleLeuProAlaLysProPheProAsnSer----- 169
OY 238 GATACACACTATTTCAGAGTTTCAATGACACTCCCTCCGTTAGACACAGACCTCT 297
Db 170 AsnSerMetTyr----- 173
OY 298 ATCTCCATTGGACACCGACCTGGAACACACAGAGTTGAAAGATGGACAAACC 357
Db 174 -----LleAspArgPro 177
OY 358 ATTCCAGAGGACGTGACAGACCAAAACATTAAAGAGATGCATCCGTAAAGAAACACAG 417
Db 178 ProSerGluLysThrProGlnGlnPro----- 186
OY 418 ATTCCTTTACACCTCTCGGCTCTCATACACTCCG-----AAGAAAG 462
Db 187 ---ProValProProGlnArgProMetAlaAlaLeuProProProProAlaGluArgAsn 205
OY 463 TACCAACCTTGCCCTCT-----GAGCCGGAGACACAGAG----- 498
Db 206 HisSerProLeuProProProGlnThrAsnHisGluGluProSerArgSerArgAsnHis 225
OY 498 ----- 498
Db 226 LysThrAlaLysLeuProAlaProSerIleAspArgSerThrLysProProLeuAspArg 245
OY 499 -----CCACCTTTATCTCAG 513
Db 246 SerLeuAlaLysProPheAspArgGluProPheThrLeuGluLysLysProProPheSerAsp 265
OY 514 AGACACACCTTTCCA-----GAAGTCCAG 537
Db 266 LysProSerIleProAlaGluArgSerLeuGluHisLeuProLysIleGluLysPro 285
OY 538 GGAATGCCACATGATAGACTTAAGGACTTAAGTGAAGTCTCTTAAGACAGAAAGTT 597
Db 286 ProLeuProProThrThrGlnArgHisGluArgSerSerProLeuProGluLysLysPro 305
OY 598 CCF-----CAT 603
Db 306 ProValProLysHisGluTyrGluProAspArgArgGluAsnAspGluAspValHis 325
OY 604 AACACAGAAAGCCTGAATCAATCATCTGTAGAAACCAAAATACT----- 651
Db 326 GlnArgProLeuProGlnProAlaLeuLeuProMetSerSerAsnThrPheProSerArg 345
OY 652 -----CAAGAGTTTCACTTGCCATTAGCACTTCTGATTCACG 690
Db 346 SerThrLysProSerProMetAsnProLeuProSerSerHisMetProGluAlaPheSer 365
OY 691 ACAAGCAACACAGTGTGCAAAAACAGATCATAGAGAGGACGACCTGTCTCTCT 750
Db 366 GluSerAsnSerPhe-----ProGlnSerAlaSerLeuProTyrPheSer 382
OY 751 CAG-----AGATGCCAGCTCCAGCCAGCTGACGCCCTACAGAAATATATAGCCCTAT 804
Db 383 GlnGluProSerAsnArgProProIleArgAlaGluGluLysArgAsnPheProLeuProLeu 402
OY 805 AAATACACAAGCTGGAGACACCTTTCCCAAAAGCTTGATAGAAAGATGTCAGAC 864
Db 403 Pro---AsnLysProArgProProSerPro-----AlaGluGluGluAsnSerLeuAsn 419
OY 865 AATGATGTGATGAGAGAAATACAGCCGACAGCAGTGAAGACCATTCATGAAGAG 924
Db 420 GluGlnThrPyrValSerTyrIleThrArgProGluAlaLeuAlaLeuArgLysIle 439
OY 925 AACAGAGATGATGTTCTTGTCGAGATGTTCCCAAAATCCAAAGAAAGACCTAT 964

```

```

Db 440 AsnGlnAspGluThrPheLeuValArgAspSerSerLysLysThrThrAsnProTyr 459
OY 985 GTTTGGCTGCTGTTTATGAAACAAAGCTTACATGTAATAATCCGTTCTGAGAGG 1044
Db 460 ValLeuMetValLeuTyrLysAspLysValTyrAsnIleGlnIleArgTyrGlnLysGlu 479
OY 1045 AATCAGCAGTTGGCCGAGGACAGCTCAGAGACATGAGAGATGAGTTCATGATGAA 1104
Db 480 SerGlnValTyrLeuLeuGluTyrGlyLeuArgGluLysGluAspPheLeuSerValSer 499
OY 1105 GACATCATCGAACACTACAG 1125
Db 500 AspIleIleAspTyrPheArg 506

RESULT 2
B56110
Lysine phosphoprotein SLP-76 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C:Accession: B56110
R:Jackson, J.K.; Morito, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretz
J. Biol. Chem. 270, 7029-7032, 1995.
A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with
A:Reference number: A56110; MIMD:95221345; PMID:7706237
A:Accession: B56110
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <JNC>
A:Cross-references: GB:U20159; NID:9806767; PIDN:AAC52189.1; PID:9806768.
C:Superfamily: SH2 homology
C:Keywords: phosphoprotein
F:422-520/Domain: SH2 homology <SH2>

Alignment Scores:
Pred. No.: 3.03e-17 Length: 533
Score: 315.50 Matches: 121
Percent Similarity: 40.94% Conservative: 44
Best Local Similarity: 30.02% Mismatches: 169
Query Match: 15.62% Indels: 69
DB: 2 Gaps: 18

US-09-856-061-3 (1-1129) x B56110 (1-533)
OY 58 GGCAGTACCAGAGATGAACAGCCCTTCTAGACTGGAGAAACTTGTCTGACATC 117
Db 126 GlyAspTyrGluSerProAsn-----GluGluGluGlnAlaLeuVal 139
OY 118 CTGATGAGACAAAGGCCACAGTATGATGATGATGATGATGATGATGATGATGATGAT 177
Db 140 AspAspAlaAla-----AspTyrGluProProProProSerAsnGlu 153
OY 178 GAGACATGGCAGTGGATTAATAATTTTACACCGCGCTATA-----AAGAAATCGAA 231
Db 154 GluAlaLeuGlnSer---SerIleLeuProProAsnSerPheHisAsnThrAsnSerMet 172
OY 232 TATGCAAGT---ACACACTATTTCAGTGTGCAATGACACTCCCTT---CCGTTAGAC 285
Db 173 TyrIleAspArgProProThrGluLysValSerGlnGlnProValProProLeuArg 192
OY 286 ACAGAGACCTCTATCTCCATTTGACAGCCAGCTGGAACACACAGACAGAGTTGGAAGA 345
Db 193 ProLysProAlaLeuPro---ProLeuProThrGluArgAsnHisSerProLeuSerPro 211
OY 346 GTGAGACAAACCATTTCCAGAGGACGTGACAGCAACCAACTTAAGCAGATGCATCCGTA 405
Db 212 ProHisProAsnHisGluGluProProSerArgSerGluAsnLys----- 226
OY 406 AGAAAGACAAGATCTCTTACCA-----CCTCTCGGCTCTC 444
Db 227 ---ThrAlaLysLeuProAlaProSerIleAspArgSerThrLysProProLeuAspArg 245
OY 445 ATACACTTCCGAAGAAGTACCAACCTTG-----CCCTCTGAGCCGGAG 489

```

Db 246 SerLeuAlaProLeuAspArgGluProPheIleLeuGlyLysLysProProPheSerAsp 265
 QY 490 AGCAGACAGCCACCTTATCTCAGACAGACACCTTCCAGAACTCCAG-----GGAATG 543
 Db 266 LysProSerAlaProLeuGlyArgGluHis---LeuProLysIleGlnLysProProLeu 284
 QY 544 CCCAGTCAGATAAGCTTAAAGGAGCTTAAGTAGGCTCTT-----GAAGCA 588
 Db 285 ProProAlaMetAspArgHisGluArgAsnGluArgLeuGlyProValThrThrArgLys 304
 QY 589 GAAAAAGTTCCTCATACCAAGAGAGAGCCGTAATCAATCATATCTGTTAGAA----- 639
 Db 305 ProSerValProArgHisGlyArgGlyProAspArgArgGluAsnAspGluAspAspVal 324
 QY 640 AACCAAAATACCTCAGAGATTCACCTGTCATAGAGATTCCTTCATTCAGACAGACAGAAC 699
 Db 325 HisGlnArgProLeuProGlnProSerLeuProSerMetSerSerAsnThrPheProSer 344
 QY 700 CACAGTGTGCA----- 711
 Db 345 ArgSerValGlnProSerSerLysAsnThrPheProLeuAlaHisMetProGlyAlaPhe 364
 QY 712 AACAGATCATATAGAGAGGACGACGACCTGTTCT-----CCTCAGACATGCCAGCCT 765
 Db 365 SerGluSerAsnIleGlyPheGlnGlnSerAlaSerLeuProProTyrPheSerGlnGly 384
 QY 766 CAGGACAGTCGACGCT-----CAGGAAATATACTGCCCTTAATATACACAGAC 816
 Db 385 ProGlyAsnArgProProLeuArgSerGluGlyArgAsnLeuProLeuProValProAsn 404
 QY 817 TGGAGACACCTTCCCAAAAGGTCTGATAGAAAGATGTCAGACCAATGAATGTATC 876
 Db 405 ---ArgProGlnProProSerProGlyGluGlnGlnThrProLeuAspGluIleTyr 423
 QY 877 ATTGGAATATACGCCGCCAGGAGAGAGAGGATTCATGAAAGGAAACAGAGATGCT 936
 Db 424 ValSerTyrIleThrArgProGluAlaGluAlaLeuArgLysIleAsnGlnAspGly 443
 QY 937 AGTTTCTGTGTCGAGATGTTCTCCACAAATTCACAGAGAGAGCCCTATGTTTGGCTGTG 996
 Db 444 ThrPheLeuValArgAspSerSerLysLysThrAlaAsnAsnProTyrValLeuMetVal 463
 QY 997 TTTTATGAGAAAGTCTACAAATGTAAATAATCGCTTCTGGAGAGAAATCAGACAGTT 1056
 Db 464 LeuTyrLysAspLysValTyrAsnIleGlnIleArgTyrGlnGlnGlnLysValTyr 483
 QY 1057 GCCCTGGGACAGAGCTCAGAGAGATGAGAGATTTGATTCAGTGAAGACATCATCGAA 1116
 Db 484 LeuLeuGlyThrGlyLeuArgGlyLysGluAspPheLeuSerValSerAspIleIleAsp 503
 QY 1117 CACTACAG 1125
 Db 504 TyrPheArg 506
 RESULT 3
 A38749
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain alpha - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Feb-1992 #sequence, revision 14-Feb-1992 #text, change 05-Nov-1999
 C:Accession: A38749
 R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, A.
 Cell 65, 91-104, 1991
 A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine ki
 A:Reference number: A38749; MUID:91191567; PMID:1707345
 A:Accession: A38749
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-724 <OTS>
 A:Cross-references: GB:M61746; GB:M61745; NID:J163476; PIDN:AAA79511.1; PID:J163477
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:333-428/Domain: SH2 homology <SH2A>

F:624-718/Domain: SH2 homology <SH2>
 Alignment Scores:
 Pred. No.: 0.000237 Length: 724
 Score: 154.50 Matches: 81
 Percent Similarity: 37.21% Conservative: 47
 Best Local Similarity: 23.55% Mismatches: 102
 Query Match: 7.65% Indels: 114
 DB: 2 Gaps: 14
 US-09-856-061-3 (1-1129) x A38749 (1-724)
 QY 406 AGAAGAACAAAGATTCCTTACCACTCTCGGCTCTCATTAACACTTCCAGAAAGTAC 465
 Db 80 LysLysIleSerProProThrProLysProArgProArgProValAlaPro 99
 QY 466 CAACCTTGGCCCTGAGCCGGAGAGACAGAG-----CCACCTTATCT 510
 Db 100 GlyProSerLysThrGluAlaAspSerGluGlnGlnAlaSerThrLeuProAspLeuAla 119
 QY 511 CAGAGACACACCTTCCAGAAAGT----- 534
 Db 120 GluGlnPheAlaProProAspValAlaProProLeuLeuIleLysLeuValGluAlaIle 139
 QY 535 -----CAGGAAATGCCAGTCAAGTAAAGCTTAAG----- 564
 Db 140 GluLysLysGlyLeuGluLysSerThrLeuTyrArgThrGlnSerSerSerAsnProAla 159
 QY 565 GACTTAAGTAGGCTCTTGAAGCAGAAAGTCTCT----- 600
 Db 160 GluLeuArgGlnLeuLeuAspCysAspThrAlaSerLeuAspLeuGluMetPheAspVal 179
 QY 601 CATTAACAGAGAGAGCCTGATCAATCAATCATCTGTAGAAACCAAAATATCAAGAGAT 660
 Db 180 HisValLeuAlaAspAlaPheLysArgTyrLeuLeuAspLeuProAsn---ProValIle 198
 QY 661 CCACCTGCCATTAAGCTTCT----- 681
 Db 199 ProValAlaValSerSerGluLeuIleSerLeuAlaProGluValGlnSerSerGlu 218
 QY 681 ----- 681
 Db 219 TyrIleGlnLeuLeuLysLysLeuIleArgSerProSerIleProHisGlnTyrPhe 238
 QY 682 -----TCATTCAGCAGACAGACACAGCT 705
 Db 239 ThrLeuGlnTyrLeuLeuLysHisPhePheLysLeuSerGlnThrSerLysAsnLeu 258
 QY 706 GTGCMAAACAGAGATCATAGAGAGAGATGCGCCCTGTTCTCCAGAGATGCCAGCT 765
 Db 259 LeuAsnAlaArgValLeuSerGluLeuPheSerProLeuLeuPheArg-PheProAla 278
 QY 766 CCAGCCAGCTGCAGCCCTCAGAAATATACTGCCCTAATAATAC-----ACAAGC 816
 Db 278 Laser-SerGluAsnThrGlnHisLeuIleLysIleIleGlnIleLeuIleSerThrGlu 297
 QY 817 TGGAGA-----CCACCTTCCCAAAAGS----- 840
 Db 298 ThrAsnGluArgGlnProAlaProAlaLeuProProLysProProLysProThrVal 317
 QY 841 -----TCTGATAGAAAGAGATGTCAGACCAATGAATGGTATCTGAGAA 885
 Db 318 AlaAsnAsnGluMetAsnAsnAsnMetSerLeuGlnAspAlaGluTyrTyrTyrGlyAsp 337
 QY 886 TACAGCCGCCAGGAGAGTGAAGAGCATTCATGAAGGAAACAGAGATGATGTTCTG 945
 Db 338 IleSerArgGluGluValAsnGluLys---LeuArgAspThrAlaAspGlyThrPheLeu 356
 QY 946 GTCCAGATTTGTCACAAATTCAGAGAGAGAGCCCTATGTTGAGCTGTTTATGAG 1005
 Db 357 ValArgAspAlaSerThrLysMetHisGlyLys---TyrThrLeuThrLeu----- 372
 QY 1006 AACCAAGTCTACAAATGTAAAA---ATCGCTTCTGAGAGAGAAATCAGACAGTTGCCCTG 1062

```

Db 373 ArglysglylylAsnAsnLysLeuIleLeuPheNisArgspLysTyrGlyPhe 392
QY 1063 GGGACAGGAGCTCAGAGAGATGAGAGATTGATTCAGTGAGACATTCAGACCTAC 1122
Db 393 SerAspProLeu-----ThrPheAsnSerValValGluLeuIleAsnHisTyr 408
QY 1123 AAGAAAT 1128
Db 409 ArgAsn 410

RESULT 4
A:38748
3-phosphatidylinositol kinase (EC 2.7.1.-) 85k chain - human
C:Species: Homo sapiens (man)
C:Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 16-Jul-1999
C:Accession: A38748; S28402
R:Skolnik, E.Y.; Margolis, B.; Mohammadi, M.; Lowenstein, E.; Fischer, R.; Drepps, A.; U
Cell 65, 83-90, 1991
A:Title: Cloning of p13 kinase-associated p85 utilizing a novel method for expression/cl
A:Accession: A38748
A:Reference number: A38748; MUID:91191565; PMID:1849461
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-724 <SR>
A:Cross-references: GB:M61906
R:Panayotou, G.; Bax, B.; Gout, I.; Federlisch, M.; Wroblewski, B.; Dhand, R.; Fry, M.J.
EMBO J. 11, 4261-4272, 1992
A:Title: Intersection of the p85 subunit of PI 3-kinase and its N-terminal SH2 domain wit
A:Reference number: S28402; MUID:93049176; PMID:1330535
A:Accession: S28402
A:Status: Preliminary
A:Molecule type: Protein
A:Residues: 301-311,315-319;424-439 <PAN>
A:Gene: GDB:PIK3R1
A:Cross-references: GDB:127604; OMIM:171833
A:Map position: 5q12-5q13
C:Superfamily: SH2 homology
C:Keywords: phosphotransferase
F:333-428/Domain: SH2 homology <SH2A>
F:624-718/Domain: SH2 homology <SH2>

Alignment Scores:
Pred. No.: 0.000496 Length: 724
Score: 150.50 Matches: 82
Percent Similarity: 37.04% Conservative: 48
Best Local Similarity: 23.36% Mismatches: 93
Query Match: 7.45% Indels: 128
DB: 2 Gaps: 17

US-09-856-061-3 (1-1129) x A38748 (1-724)
QY 406 AGAAGACAGAGATTCCTTACACCTCGGCTCATACACTTCGAGAGAGATAC 465
Db 80 LysLysIleSerProThrProLysProArgPro-----ProArg----- 93
QY 466 CAACCTTGGCCCTGACCGGAGAGACAGCAG----- 498
Db 94 ---ProLeuProValAlaProGlySerSerLysThrGluAlaAspValGluGlnAla 112
QY 499 -----CCACCTTATCTTCAGAGACACACCTTCCAGAGATC----- 534
Db 113 LeuThrLeuProAspLeuAlaGlnPheAlaProAspIleAlaProProLeuLeu 132
QY 535 -----CAGGGAATG----- 543
Db 133 IleIleLeuValGluAlaIleGluLysIleGluLysSerThrLeuTyrArgThr 152
QY 544 CCCAGCTAGATTAAGGAGACTTAAGTGAAGTCTTGAAGCAGAAAAGTTCCTCAT 603
Db 153 GlnSerSerSerAsnLeuAlaGluLeuArgGlnLeuLeuAspCysAsp---ThrProSer 171

```

```

QY 604 AACGAGAGAGAGCTGATCACTCATCTGTAGAAACCA----- 645
Db 172 ValAspLeuGluMetIleAspValHisValIleAlaAspAlaPheLysArgTyrLeuLeu 191
QY 646 -----AATACTCAAGAGATTCCTGCTTGCATTAAGCTCT----- 681
Db 192 AspLeuProAsnProValIleProAlaIleValTyrSerGluMetIleSerLeuAlaPro 211
QY 681 ----- 681
Db 212 GluValGlnSerSerGluGluTyrIleGlnLeuLeuLysLeuLeuIleArgSerProSer 231
QY 682 -----TCA 684
Db 232 IleProHisGlnTyrTrpLeuThrLeuGlnTyrLeuLeuLysHisPhePheLysLeuSer 251
QY 685 TTCACGACAGACAGAACAGCTGCGAAAGAGATCATAGAGAGCGATGAGCCCTGT 744
Db 252 GlnThrSerSerLysAsnLeuLeuAsnAlaArg-ValLeuSerGluIlePheSerProk 271
QY 745 TCTCCCTCAGA-----GATGCCAGCCTCCAGCAGCTGACGCCCTCAGCAAA---ATATA 795
Db 271 LeuPheArgPheSerAlaAlaSerSerAspAsnThrGluAsnLeuIleLysValIleG 291
QY 796 CTGCCCTATTAATACACAGCTGAGAG-----CCACCTTTC 831
Db 291 uIleLeuIleSerThrGln--TrpAsnGluArgGlnProAlaProAlaLeuProProLys 310
QY 832 CCCAAAGC-----TCTGATAGAAAGGATGTCACAGC 864
Db 311 ProProLysProThrThrValAlaAsnAsnGlyMetAsnAsnMetSerLeuGlnAsn 330
QY 865 AATGATGTGATCATGAGATACAGCCGCGAGAGAGAGAGGAGCATTCGAAAGAG 924
Db 331 AlaGluTrpTyrTrpGlyAspIleSerAlaGluGluValAsnGlnLys---LeuArgSp 349
QY 925 AACAGAGATGAGTGTCTTGTGCGAGATTGTCCACAAATCCAGAGAGAGCCCTAT 984
Db 350 ThrAlaAspGlyThrPheLeuValAlaArgAspAlaSerThrLysMetHisGlyAsp---Tyr 368
QY 985 GTTTGGCGTGTGTTTATGAGAACAAAGTCAATGTAAAA---ATCCGCTTCCTGGAG 1041
Db 369 ThrLeuThrLeu-----ArgLysGlyLysAsnLysLeuIleLysIlePheHis 385
QY 1042 AGAATCAGACAGATTCCTGCGGAGACAGACTCAGAGAGAGATGAGAGATTCATCAGTA 1101
Db 386 ArgAspArgLysTyrGlyPheSerAspProLeu-----ThrPheSerSerVal 401
QY 1102 GAGACATCATCAGAACACTACAGAAAT 1128
Db 402 ValGluLeuIleAsnHisTyrArgAsn 410

RESULT 5
A:38747
3-phosphatidylinositol 3-kinase (EC 2.7.1.-) 85k chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 16-Jul-1999
C:Accession: A38747
R:Escobedo, J.A.; Navarokasatusas, S.; Kavanagh, W.M.; Milfey, D.; Fried, V.A.; Will
Cell 65, 75-82, 1991
A:Title: cDNA cloning of a novel 85 kd protein that has SH2 domains and regulates bin
A:Reference number: A38747; MUID:91191564; PMID:1849460
A:Accession: A38747
A:Molecule type: mRNA
A:Residues: 1-724 <ESC>
A:Cross-references: GB:M60651
C:Comment: This protein binds a phosphotyrosine-containing sequence of ligand-activat
3-phosphatidylinositol at position 3 of the inositol ring.
C:Superfamily: SH2 homology
C:Keywords: phosphotransferase
F:333-428/Domain: SH2 homology <SH2A>
F:624-718/Domain: SH2 homology <SH2>

```


Alignment Scores:

Pred. No.: 0.000654 Length: 724
 Score: 149.00 Matches: 76
 Percent Similarity: 32.76% Conservative: 39
 Best Local Similarity: 21.65% Mismatches: 96
 Query Match: 7.38% Indels: 140
 DB: 2 Gaps: 17

US-09-856-061-3 (1-1129) x A38747 (1-724)

```

OY 421 CCTTACACACCTCCCTGCGCTCTCATTAACACTCCGAGAGAGTACCAACCTTGCCCTC 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 ProthrProlyspProthrPro-----ProArg-----ProleuProVal 97
OY 481 GAGCGGAGAGCAGCAGG-----CCACT 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 AlaProGlySerSerLysThrGluAlaAspThrGluGlnGlnAlaLeuProLeuProAsp 117
OY 505 TTATCTCAGAGACACACCTTTCCAGAGAGTCCAGAGATGCCAGTCAGATTAAG 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 LeuAlaGluGlnPheAlaProProAspVal---AlaProProLeuLeuLeuLeuLeu 136
OY 565 GACTTAAGTGAAGTC-----CTTGAAGCAGAAAAAGTTCTCATTAACCAAGAGAGCCT 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 GluAlaIleGluLysGlyLysGlyLysSerThrLeuTyrArgThrGlnSerSer 156
OY 619 GAATCACTCATCTCTTGA---AACCAAAATCTCAAG 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 AsnProAlaGluLeuArgGlnLeuLeuAspCysAspAlaAlaSerValAspLeuGluMet 176
OY 658 ATTCACCTGGCATTAGAGATCTTCATTCATCAGACAGACAGCAAGTCGCAAAACAGA 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 IleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuAlaAspLeuProAsn--- 195
OY 718 GATCATAGAGAGGAGCATGACGCCCTGTTCTCTCT----- 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 -----ProValIleProValAlaValTyrAsnGluMetMetSer 208
OY 751 -----CAGAGATGCCAGGCTCCAGCCAGCTGC----- 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 LeuAlaGlnGluLeuGlnSerProGluAspCysIleGlnLeuLeuLysLysLeuIleArg 228
OY 778 -----AGCCCTCAGCAAAATATATCTGCCCTATTAATAC----- 810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 LeuProAsnIleProHisGlnCysTrpLeuThrLeuGlnTyrLeuLeuLysHisPhePhe 248
OY 810 ----- 810
Db 249 LysLeuSerGlnAlaSerSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePhe 268
OY 810 ----- 810
Db 269 SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGluHisLeuIleLys 288
OY 811 -----ACAAGCTGG----- 819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 AlaIleGluIleLeuIleSerThrGluTrpAsnGluArgGlnProAlaProAlaLeuPro 308
OY 820 -----AACCACCTTCCCTC-----AAAAGCTCGATGAAAGAGATGTC 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 ProLysProProLysProThrThrValAlaAsnAsnSerMetAspAsnAsnMetSerLeu 328
OY 859 CAGCACAATGAATGATACATGAGAAATACAGCCGACGAGTGAAGAGCATTCAG 918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 GlnAspAlaGluTrpTyrTrpLysAspIleSerArgGluGluValAsnGluLys---Leu 347
OY 919 AAGGAGAACAGAGATGATGTTCTGTCGAGATGTTGCCAACAATCCAAGAGAG 978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp 367
OY 979 CCTTATGTTTGGCTGTGTTTATGAGAACAAAGTCTACAATGTAA---ATCCGCTTC 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 368 ---TyrThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIle 383
  
```

```

OY 1036 CTGAGAGCATCAGACGTTGCTCCCTGGGAGACAGCATCAGAGATGAGAAGTTGAT 1095
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 PheHisArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheAsn 399
OY 1096 TCAGTAGAAGACATCATCAGACCTACAGAAAT 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 SerValValGluLeuIleAsnHisTyrArgAsn 410
  
```

RESULT 6

149553
 N:Protein-tyrosine kinase (EC 2.7.1.112) emb - mouse
 N:Alternate names: B-cell progenitor cytoplasmic tyrosine kinase; Bruton agammaglobulin
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Feb-2000
 C:Accession: I49553; B45184; J004771; PC4366
 R:Slideras, P.; Muller, S.; Shlets, H.; Jin, H.; Khan, W.N.; Nilsson, L.; Parkinson, E
 ulos, K.G.

J. Immunol. 153, 5607-5617, 1994
 A>Title: Genomic organization of mouse and human Bruton's agammaglobulinemia tyrosine
 A:Reference number: I49553; MUID:95081608; PMID:7989760

A:Accession: I49553
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA

A:Residues: 1-659 <SID>
 A:Cross-references: GB:L29788; NID:9625143; PIDN:AA66943.1; PID:9625144
 R:Tsukada, S.; Saffran, D.C.; Rawlings, D.J.; Parolimi, O.; Allen, R.C.; Klisak, I.;
 Cell 72, 279-290, 1993

A>Title: Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-link
 A:Reference number: AA5184; MUID:93145329; PMID:8425221
 A:Accession: B45184

A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA

A:Residues: 1-122, P, 124-659 <TSU>
 A:Cross-references: GB:L08967; NID:9192233; PIDN:AA37316.1; PID:9192234
 A:Experimental source: 702/3 pre-B cell line

A>Note: sequence extracted from NCBI backbone (NCBIP:123834) and corrected to corresp
 R:Yamada, N.; Kawakami, Y.; Kimura, H.; Fukumachi, H.; Balier, G.; Altman, A.; Kato, T
 Biochem. Biophys. Res. Commun. 192, 231-240, 1993

A>Title: Structure and expression of novel protein-tyrosine kinases, Emb and Ent, in
 A:Reference number: J00471; MUID:9323578; PMID:8476425
 A:Accession: J00471

A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA

A:Residues: 1-66, 'A', 68-197, 'W', 198-449, 451-659 <YAM>
 A:Cross-references: GB:L10627; NID:9193016
 A>Note: this translation is not annotated in Genbank entry MUSEMBX, release 116.0

R:Kojima, T.; Fukuda, M.; Watanabe, Y.; Hamazato, F.; Mikoshiba, K.
 Biochem. Biophys. Res. Commun. 236, 333-339, 1997
 A>Title: Characterization of the pleckstrin homology domain of Btk as an inositol pol

A:Reference number: PC4366; MUID:97382431; PMID:9240435
 A:Accession: PC4366
 A:Molecule type: mRNA

A:Residues: 1-165 <KO>
 C:Comment: This protein specifically recognizes the inositol 1,3,4,5-pentakisphosphat
 e molecule to the membrane.

C:Genetics:
 A:Gene: emb; Btk
 C:Superfamily: protein-tyrosine kinase tec; pleckstrin repeat homology; protein kinas
 C:Keywords: ATP; phosphoprotein; phosphotransferase; tyrosine-specific protein kinase
 F:2-131/Domain: pleckstrin repeat homology <PIK>

F:221-269/Domain: SH3 homology <SH3>
 F:281-377/Domain: SH2 homology <SH2>
 F:400-658/Domain: protein kinase homology <KIN>

F:408-416/Region: protein kinase ATP-binding motif
 F:551/Binding site: phosphate (Tyr) (covalent) #status predicted

Alignment Scores:

Pred. No.: 0.00195 Length: 659
 Score: 143.00 Matches: 60
 Percent Similarity: 33.98% Conservative: 27
 Best Local Similarity: 23.44% Mismatches: 71
 Query Match: 7.08% Indels: 98

Dn - 375 ProValProValAlaIalaProArgProProValGlnInLysMetProProValaThr 394
QY 817 TGGAGACCACCTTTCCCAAAAGGTCTGAATAGAAGAGATTCACGCAC----- 864
Db 395 ValArgProAlaValLeuProArgProGluAsnThrProLeuProHisLeuGlnArgSer 414
QY 864 ----- 864
Db 415 ProProaspGlyInserPheArgGlyPheSerPheGluLysAlaArgInProSerGln 434
QY 865 -----AATGAA 870
Db 435 AlaAspThrGlyGluGluAspSerAspGluAspTyrGluLysValProLeuProAsnSer 454
QY 871 TGGTACATTGGAGATACAGCCCGCAGGACAGTGGAAGAGCGCATTCATGAAG----- 921
Db 455 ValPheValAsnThrInglInUserCySgLuValGluInArgLeuPheLysAlaThrAspPro 474
QY 922 ----GAGAACAGATGGTAGTTTCTGGTCCGAGATGTTGCCACAAAATCCAGGAA 975
Db 475 ArgGlyGluProGlnAspGlyLeuTyrCysIleArgAsnSerThrLysSerGlyGly 494
QY 976 GAGCGCATGTTTGGCGTGTATTATGAG-----AACAAGCTCAACATGTAATAATC 1029
Db 495 -----ValLeuValValTyrAspSerLysSerAsnLysValaArgAsnTyrArgIle 511
QY 1030 CGCTTCTGGAGAGAAATCACAGACTTGGCCCTGGGAGCAGACGATCAGAGATGAGAG 1089
Db 512 -----PheGluLysAspSerLysPheTyr-----LeuGluGlyGluValLeu 525
QY 1090 TTTCATTGATAGAGAACATCATCATCGAACACTAC 1122
Db 526 PheAlaSerValIGlySerMetValIGluHisTyr 536

RESULT 8
A45184
N:protein-tyrosine kinase (EC 2.7.1.112), nonreceptor type, BTK - human
A:Alternate names: atk; B cell progenitor kinase; Bruton agammaglobulinemia tyrosine kin
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence.revision 02-Jul-1996 #text_change 21-Jul-2000
C:Accession: I37212; I54541; I39457; S28912; A45184
R:Ohno, Y.; Haide, R.N.; Litman, R.T.; Fu, S.M.; Nelson, R.P.; Kratz, J.; Kornfeld, S.J.
Proc. Natl. Acad. Sci. U.S.A. 91, 9067-9066, 1994
A>Title: Genomic organization and structure of Bruton agammaglobulinemia tyrosine kinase
mmaglobulinemia.
A:Reference number: I37212; MUID:94377492; PMID:8090769
A:Accession: I37212
A>Status: nucleic acid sequence not shown; translation not shown; not compared with con
A:Molecule type: DNA
A:Residues: 1-659 <OHT>
A:Cross-references: EMBL:U10087; NID:g1226232; PIDN:AAB60639.1; PID:g517438
A>Note: only intron-exon junctions are shown
R:Roher, J.; Parolin, O.; Belmont, J.W.; Conley, M.E.
Immunogenetics 40, 319-324, 1994
A>Title: The genomic structure of human BTK, the defective gene in X-linked agammaglobuli
A:Accession: I54541
A>Status: nucleic acid sequence not shown; translated from GB/EMBL/DBDUT
A:Molecule type: DNA
A:Residues: 1-659 <ROH>
A:Cross-references: GB:IJ1572; NID:g632958; PIDN:AAA61479.1; PID:g632960
R:Hagemann, T.L.; Chen, Y.; Rosen, F.S.; Kwan, S.P.
Hum. Mol. Genet. 3, 1743-1749, 1994
A>Title: Genomic organization of the Btk gene and exon scanning for mutations in patient
A:Reference number: I39457; MUID:95152493; PMID:7860320
A:Accession: I39457
A>Status: Preliminary; translated from GB/EMBL/DBDUT
A:Molecule type: DNA
A:Residues: 1-278,281-659 <HAG>
A:Cross-references: EMBL:U13433; NID:g575888; PIDN:AA651347.1; PID:g575890
R:Verdie, D.; Vorchovsky, I.; Sideras, P.; Holland, J.; Davies, A.; Flinter, F.; Hammar
Nature 361, 226-233, 1993

A>Title: The gene involved in X-linked adammaglobulinemia is a member of the src fam
A:Reference number: S28912; MUID:93140868; PMID:8380905

A:Accession: S28912

A:Molecule type: mRNA

A:Residues: 1-659 <VERT>

R:Cross-references: GB:558957; NID:q312466; PIDN:CAA41728.1; PID:q312467; GB:U78027;
R:Tsuchida, S.; Saffran, D.C.; Rawlings, D.J.; Petrolini, O.; Allen, R.C.; Klsak, I.;
Cell 72, 279-290, 1993

A>Title: Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-link
A:Reference number: A45184; MUID:93145329; PMID:8425221

A:Accession: A45184

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-19, 'R', 21-190, 'A', 192-206, 'T', 208-210, 'I', 212-218, 'E', 220-238, 'E', 240-2

A:Cross-references: GB:U78027; GB:L35265; NID:g2281904

A:Experimental source: erythroleukemia cell line K562

A>Note: sequence extracted from NCBI backbone (NCBIF:123835)

C:Genetics:

A:Gene: GDB:BTK; AGMX1, IMD1

A:Cross-references: GDB:120542; OMIM:300300

A:Map position: Xq21.33-qx22

A:Introns: 47/3; 80/3; 103/3; 131/1; 174/1; 196/3; 259/2; 280/2; 298/3; 325/2; 368/1;::

C:Superfamily: ATP, phosphotransferase; tyrosine-specific protein kinase

F:2-131/Domain: pleckstrin repeat homology <PLK>

F:221-269/Domain: SH3 homology <SH3>

F:281-377/Domain: SH2 homology <SH2>

F:408-658/Domain: protein kinase homology <KIN>

F:408-416/Region: protein kinase ATP-binding motif

Alignment Scores:

Pred. No.:	Length:	Matches:
Score: 0.00339	659	58
Percent Similarity: 140.00	Conservative: 29	Mismatches: 71
Best Local Similarity: 33.98%	Indels: 98	Gaps: 11
Query Match: 22.66%		
DB: 2		

US-09-856-061-3 (1-1129) x A45184 (1-659)

Oy 373 AGAACGCAAAACATTAAAGAGATGCATCCGTGAAGAACAAGATCTTTACCACCT 432

Db 171 ATgaaNglySerLeuLysProGlySerSerHisArgLySthrLysLysProLeuProPro 190

Oy 433 CCTCGCCTTCATTAACACTTCCGAAGAATCCAAACCTTGCCCCGAGCCGGAGAGC 492

Db 191 ThrProGluGlnAspGlnHleuLeuLys-----ProlenProProGluProAlaIla 208

Oy 493 AGCAGGCCACTTTCATCTCACAGACACACTTTCCAGAAATCCAGGAGATGCCCATCTGAG 552

Db 209 -----AlaPro 210

Oy 553 ATAAGCTTAAGGAGCTTAAGTAGSGCTGTGAAGA---GAAAAGTTCCTCATTAACAG 609

Db 211 ValSerThrSerGlnLeuLysValValAlaLeuTYrAsPTyrMetProMetasnIa 230

Oy 610 -----AGNAAGCTGAATCAATCACTCATCTGTGAAGAACCAAAATACTCAA 654

Db 231 AsnaSplenuGlnLeuArgLysLysaspGluYrPhelHleuGlnGlnSerasn----- 248

Oy 655 GAGATTCACACTTGCCATTAGCAGTTCTTCATTCACGACAGCAACACACAGTGTGCAAAAC 714

Db 249 ---LeuProTrp-----TrpArgIa 254

Oy 715 AGAGATCATAGAGGAGCATGCAAGCCCTGTCTCCTCAGAGATGCGACGCTCCACAGCAGC 774

Db 255 ArgasplYsasngly----- 259

Oy 775 TSCAGCCCTCAGCAAAATATACTGCCCTATAATACACAGCGTGAGACACACTTTCCCC 834

Db 260 -----GlnGlnGlyTYrIleProSerasnTYrValTh----- 270

Oy 835 AAAAGGTCTGATAGAAAGGATGTCAGCACCAATGAATGGTATCTTGGAATATACAGCCGC 894

Db 271 -----GluAlaGluAspSerIleGluMetGlyGluTrpTyrSerLysHisMetThrArg 288
 Oy 895 CAGCAGATGGAGGACATTCATGTAAGAGAACAGATGCAAGTTCTGTGTCGAGAT 954
 Db 269 SerGlnAlaGluGlnLeuLeuLysGlnGlyLysGlnGlyGlyPheIleValArgAsp 308
 Oy 955 TGTTCACAAAATCCAGAGAACCCCTATGTTTGGCTGTGTTTATGAGAACAAAGTC 1014
 Db 309 SerSerLysAlaGlyLys----- 314
 Oy 1015 TACATGTGTAATAATCCGCTTCCTGGAGAGAAATCAGCAAGTTGCCCTGGCAGCACTC 1074
 Db 315 TyrThrValSerVal-----PheAlaLysSerThrGly--- 325
 Oy 1075 AGAGAGATGAGAACTGTTGATCTAGTAAGACATCATCGAACAACATAC 1122
 Db 326 -----AspProGlnGlyValIleArgHisTyr 334
 RESULT 9
 TVMVR
 protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rashee
 C:Species: feline sarcoma virus
 A:Note: host Felis sp. (cat)
 C:Date: 27-Nov-1985 #sequence_revision 26-May-1995 #text_change 31-Mar-2000
 C:Accession: A00653; A03937
 R:Naharro, G.; Robbins, K.C.; Reddy, E.P.
 Science 223, 63-66, 1984
 A:Title: Gene product of v-fgr onc. hybrid protein containing a portion of actin and a
 A:Reference number: A00653; MUID:84097512; PMID:6318314
 A:Accession: A00653
 A:Molecule type: DNA
 A:Residues: 1-663 <NAH>
 A:Cross-references: GB:X00255; GB:K01487; NID:961542; PID:CAA25063.1; PID:g61543
 A:Note: the authors translated the codon GAT for residue 14 as Glu
 C:Comment: This protein is synthesized as a gag-fgr polypeptide.
 C:Genetics:
 A:Gene: fgr
 C:Superfamily: feline sarcoma virus protein-tyrosine kinase fgr; protein kinase homology
 C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; poly
 F:1-118/Region: gag polypeptide similarity
 F:141-268/Region: actin similarity
 F:285-382/Domain: SH2 homology <SH2>
 F:402-660/Domain: protein kinase homology <KIN>
 F:410-418/Region: protein kinase ATP-binding motif
 F:432/Active site: Lys #status predicted

Alignment Scores:

Prod. No.: 0.00408 Length: 663
 Score: 139.00 Matches: 83
 Percent Similarity: 35.14% Conservative: 60
 Best Local Similarity: 20.39% Mismatches: 137
 Query Match: 6.88% Indels: 127
 DB: 1 Gaps: 17

US-09-856-061-3 (1-1129) x TVMVR (1-663)

Oy 95 GGAAGAGAACTTCTGTCAGTCTGATGAG-----CAAAAGCCACAGTATGATG 148
 Db 27 GAlaAlaGluValArgLysLysTrpIleThrLeuGlnGlnAlaGluTrpValMetMet 46
 Oy 149 ACTTGATGACCCCTGACCTTGGATGCAAGACATGGCAGCATTAATAATTTTACAG 208
 Db 47 AsnVal-----GlyTrpProArgGlnGlyThr----- 55
 Oy 209 CCCGCGCATTAAGAAATCTGAATATGACATACACTATTTCAAGTTGCATGAC 268
 Db 56 -----PheSerLeuAspAsnIleSerGlnValGlnLys 67
 Oy 269 CTCCTCTTCCTAGACACAGCAGCATCTATCTCATTTGACACCGCAGTGAACAC 328
 Db 68 Ile-----PheAlaProGlyProTyrGlnHisProAspGlnValPro-----Tyr 82

Oy 329 AGACAGCTTGAAA-----GAGTGACAAACCATTTCCAGG 367
 Db 83 IleThrThrTrpArgSerLeuAlaThrAspProProSerTrpValArgPro----- 99
 Oy 368 ACGTCAGAAACCCAAAACATTAAGAGATGCATCCGTAAAGAAAGAAAGATTCCTTAC 427
 Db 100 -----Phe-LeuP 102
 Oy 428 CACCTCTCCGCTCTCATTAACACTTCGGAAGAAATACACCCCTTGGCC----- 477
 Db 102 ToProProLysProProThrSerLeuProGlnProHisSerProGlnProAlaArgAla 122
 Oy 478 -----CCTGAGCCGAGAGACAGCAGCCACTTATCTCAG- 513
 Db 122 euCyArgProAlaValArgArgProAlaProLeuProProLeuProProThrAlaMetG 142
 Oy 513 ----- 513
 Db 142 LuGluGluValAlaIleValIleAspAsnGlySerGlyMetCysLysAlaGlyPhe 162
 Oy 514 -----AGACACACCTTCCAGAACTCCAGGAAATGCCACTCAG----- 552
 Db 162 IaGlyAspAspAlaProArgAlaValPheProSerIleValGlyArgProArgHisGln 182
 Oy 553 -----ATAAGCTTAAGGAC-----TTAAGAGAGCTCTGAGCAGAAA 592
 Db 182 LyValMetValGlyMetGlyGlnLysAspSerTyrValGlyAspGlnAlaGlnSerLys 202
 Oy 593 AAGTTCCTCATTAACAGAGAAACCTGAAATCAACTCTGTAGAAACCAAAATACTC 652
 Db 202 rgGlyIleLeuThrLeuLysTyrProIleGlnHisGlyIleValThrAsnTrpAspAsm 222
 Oy 653 AAGAAATTCACCTTCCATTTAGCACTTTCATTCACAGCAAGCAGCAGTGTGCAA 712
 Db 222 etGlu-----LysIleThrHisThrPheTyr 232
 Oy 713 ACAGAGATC---ATGAGAGAGCATGACCCCTGTCTCTCAGAGATGCCAGCTCCAG 769
 Db 232 sn-GluLeuArgValAlaProGlnGlnHisProValIleLeuThrGlnAlaProLeuAsn 251
 Oy 770 CCAGCTGACGCCCTCAGCAAAATATCTGCCCTTAATAATACAAAGCTGAGACCCCTT 829
 Db 252 ProLysAla--AsnArgGlnLysMetThrGlnIleMetPheGluThrPheAsnIlePro 271
 Oy 830 TCCCAAAAGGTGATGAGAAAGATGTCAGCAACATGATGATGATGAGATGCA 889
 Db 271 eraThrValAlaProValAspSerIleGlnAlaGluGluTrpArgHisGlyLysIleG 291
 Oy 890 GCCGCGCAGCAGTGAAGAGCATTCATGAAAGAG---AACAGATGATGATTTCTTG 946
 Db 291 LyArgLysAspAlaGlnArgGlnLeuLeuSerProGlyAsnAlaArgGlnAlaPheLeu 311
 Oy 947 TCCGAGATGTTCCAAAATCCAAAGAGACCCCTATGTTTGGCTGTG----- 966
 Db 311 alArgGlnSerGluThr-----ThrLysGlyAlaTyrSerLeuSerIleArgAspTrp 329
 Oy 997 -----TTTATGAGAAACAAAGCTCAATGATTAATAATCCGCTCTCGAGAGATCAGC 1051
 Db 329 spGluAlaArgGlyAspHisValLysHisTyrLysIleArgLysLeuSphrGly---- 347
 Oy 1052 AGTTTCCCTGGGAGCAGCTCAGAGAGATGAGAAAGTTGATTCATGACATGACATCA 1111
 Db 348 -----GlyTyrTyrIleThrThrArgAlaGlnPheAsnSerValGlnLeu 364
 Oy 1112 TCGAACAATAC 1122
 Db 364 alGlnHisTyr 367
 RESULT 10
 FOMVGM
 gag-abl polypeptide - Abelson murine leukemia virus
 N:Contains: amino end of core shell protein p30; core protein p15; inner coat protein
 C:Species: Abelson murine leukemia virus

Db 50 GluValThrHis---IleLysIleGlnAsnGlyAspPheAspLeuTyGlyGly 68
 QY 679 TCCTCATTCAGACACCAACAGAGTGTGCAAAACAGACATCTATACAGAGCATCCAG 738
 Db 69 GluLysPheLathrLeuProGluLeuValGlnTyTrMetGluAsnGlyGluLeuLys 88
 QY 739 CCCTGTCTCTCTAG-----AGATGCCAGCTCCAGCCAGCTGACGCTCCAGCAAGAAAT 792
 Db 89 GluLysAsnGlyGlnAlaIleGluLeuLysGlnProLeuIleCysAla----- 104
 QY 793 ATACTGCTTATTAATACACAGCTGGAGACACACCTTCCCAAAAGCTGTGATAGAAAG 852
 Db 105 -----GluProThrThrGluArg----- 110
 QY 853 GATGTCCAGCAATGATGATCATGTGAGAAATACAGCCGCGCGAGAGTGAAGAGCA 912
 Db 111 -----TrpPheAsnGlyAsnLeuSerGlyLysGluAlaGluLysLeu 124
 QY 913 TTCATGAGAGAGAACAGATGATGATGCTGTGCTCCGAGATGTTCACAAATCCAG 972
 Db 125 IleuGluLysGlyLysAsnGlySerPheLeuValArgLys-----SerGlnSerLys 142
 QY 973 GAAGAGCCCATATGTTGGCTGTGTTTATGACAAAGCTCAATGTAATAATCCGC 1032
 Db 143 ProGluAspPheValLeuSerValArgThrAspAspLysValThrHisValMetIleArg 162
 QY 1033 TTCCTGGAGAGATGAGATGCTTGGCTGGGAGAGACTCAGAGAGATGAGAACTT 1092
 Db 163 Trp-----GlnAspLysLysTyArgAspValGly-----GlyGlyGluSerPhe 176
 QY 1093 GATTCACTAGACAGATCATGCAACACTAGACAG 1125
 Db 177 GlyThrLeuSerGluLeuIleAspHisLysTyLys 187
 RESULT 12
 T37550
 hypothetical coiled-coil protein [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 08-Sep-2000
 C:Accession: T37550
 R:Conor, R.; Churcher, C.M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: Z21723
 A:Accession: T37550
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-485 <CON>
 A:Cross-references: EMBL:AL132714; PTDN:CA859797.1; GSPDB:GN00066; SPDB:SPAC11H1.01
 C:Experimental source: strain 972h-; cosmid c11H11
 C:Genetics:
 A:Gene: SPDB:SPAC11H1.01
 A:Map position: 1
 Alignment Scores:
 Pred. No.: 0.0157 Length: 485
 Score: 131.50 Matches: 74
 Percent Similarity: 40.41% Conservative: 65
 Best Local Similarity: 21.51% Mismatches: 118
 Query Match: 6.51% Indels: 87
 DB: 2 Gaps: 18
 US-09-856-061-3 (1-1129) x T37550 (1-485)
 QY 163 GAGCTTCGATGAGAGACATGCGACGATTAATTTTACACAGCCGCGCTATGAAG 222
 Db 154 LysLeuLysValLysGluGluAsnGluAlaProProValIleProAlaLysProPheSer 173
 QY 223 GAATTCGATATATCCAGATACACACTATTTCAAGTT---GCAATGGACACT----- 270
 Db 174 SerSer-----SerGluGlnHisPheArgLysValProAlaLeuProSerLysLeuPro 191
 QY 271 CCCCTTCCTTATACACACAGGACCTCTATGTCATTTGAGACGCCGAGCTTGAACACACAG 330
 Db 271 CCGCTTCCTTATACACACAGGACCTCTATGTCATTTGAGACGCCGAGCTTGAACACACAG 330

Db 192 ProLysProLeuLysIleThrAlaAsnSerSerLeuGlyGlnGluThrAsnSerAsnSer 211
 QY 331 ACGAGGTTGGAAAGAGG-----GACAAACCCATTTCAGAGGAGCTCAGA 375
 Db 212 SerSerPheGlnSerThrLeuPheSerLeuAsnThrAlaProPheSer-----AlaThr 229
 QY 376 AGCCAAACATTTAAAGAGATGCA---TCCGTAAAGAAAGACAGATTCCTTACCACT 432
 Db 230 SerGlnGlnLeuValHisAspSerValSerLeuArg----- 241
 QY 433 CCTGCGCTCTCATACACTTCCAGAGAGTACCAACCCCTTGCCCTGACCGCGAGAGC 492
 Db 242 ---ArgProSerSerAsnIleProAlaGln---LysProIleProProLysProGluGln 259
 QY 493 AGCAGG-----CCACCTTATCTCGAAGACACACACTTCCAGAA 531
 Db 260 AsnGluIleIleIleThrLysAspThrProSerLeuLysAspLysTySerLysProAla 279
 QY 532 GTCCAGGGAATGCCAGTCAATAGCTTAAGGACTTAAGTACAGTCTTGAAGAGAA 591
 Db 280 LeuLeuProGlnLysProLysValSer-----LysGlyGlnIleValGlnGlnVal 296
 QY 592 AAATCTCTCATAAACAGAGAGAGCTGAAATCAATCATCATCTGTTGAAGAAACAAATACT 651
 Db 297 SerValPheSerThrGlyLysLysIleGluSerGlnSerLeuLeuAsnLeuIleAspThr 316
 QY 652 CAA---GAGATTCACCTTCGATTAAGCATGCTTCTTCAATTCACGAAACCAACCAAGTGTG 708
 Db 317 AspIleGluThrProLeuLysGlyLysSer-----GluLeuLeu 329
 QY 709 CAACACAGACATCATAGAGAGAGGAGCCAGCCCTGTCTCTCAGAGATGCCAGCTTCA 768
 Db 330 TyrSerGluAspPheLysProAsnValAspProValLysIleGlnGlnIle----- 346
 QY 769 GCCAGCTGCAGCCCTCCAGAAATATATCTGCCCTATTAATACACAACTGG----- 819
 Db 347 -----LeuHisLysGlnAsnLysIleIleGluGluLysTrpIleSerGlnIleArgIle 364
 QY 820 -----AGACCACTTCCCAAAAGGCT 843
 Db 365 SerLysAsnLeuGluValLysGlnArgLeuLeuAspGlnGluArgHisAlaLeuGluThr 384
 QY 844 GATGAAAGATGTCCAGCACATGAATGGTACATTGAGAGATACAGCCGCGAGCAGTGTG 903
 Db 385 LeuValLysAsnIleGluAsnAsnArgPheIleLeuGlyLysArgValGlyAla--- 403
 QY 904 GAAGAGCATTCATGAGAGAGAACAGAGATGCTGTTCTTGGTCCGAGATTTGTCACA 963
 Db 404 ArgGluAlaLeuGlnLysLeuAspAsn-----LeuLysAspLeuSerVal 418
 QY 964 AAATCCAAGAGAGAGCCCTATGTTTGGCTGTGTTTATGAGAACAAAGTCAATGTA 1023
 Db 419 -----GlnGluLeuPheIleIleIleProSerGluArgLeuLysTyTrpGluLeu 435
 QY 1024 AAATCCGCTTCTCGAGAGAGATCAGCACTTGGCTGGGAGACAGACTCAGAGAGAT 1083
 Db 436 Lys-----ArgLysAsp 439
 QY 1084 GAGAGATTGAT 1095
 Db 440 GluLysLeuAsp 443
 RESULT 13
 B38749
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85k chain B - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 16-Jul-1999
 C:Accession: B38749
 R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, J.
 Cell 65, 91-104, 1991
 A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine
 A:Reference number: A38749; M0ID:91191567; PMID:1707345

A:Accession: B38749
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-723 <OTS>
 A:Cross-references: GB:M61745; GB:M61746
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:325-420/Domain: SH2 homology <SH2A>
 F:617-706/Domain: SH2 homology <SH2>

Alignment Scores:

Pred. No.:	0.018	Length:	723
Score:	131.00	Matches:	81
Percent Similarity:	34.75%	Conservative:	50
Best Local Similarity:	21.49%	Mismatches:	118
Query Match:	6.49%	Indels:	128
DB:	2	Gaps:	18

US-09-856-061-3 (1-1129) x B38749 (1-723)

```

QY 301 TCATTTGGA---CAGCGACCTGGACACACAGACGAGTTGGAGAGAGTGCACAAACCC 357
    |||:||||| ||| ||| :||||||| :||| ||| |||
DB 53 SerValGlyTyrMetProGlyLeuAsnGluArgThrArg---GlnArgGlyAspPhePro 71
    |||
QY 358 ATTCGAGGAGCGTCGAAAGCCAAACATTAAAGAGATGCATCCCTAAGAAAGAACAG 417
    |||
DB 72 GlyThr-----TyrValGluPheLeuGlyProValAlaLeuLeuAlaArg----- 85
    |||
QY 418 ATTCCCTTACACACTCTCGGCTCTCATTAACACTTCGAGAGATGACCAACCTTGGCC 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86 ---ProGlyProArgProArgProGlyProArgProLeuProAlaArgProArgAspGlyPro 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 CCTGACCCGAGAGCAGCAGCGACCTTATGTCAGACACACCTTCAGAGATGACAG 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 105 ProGluProGlyLeuThrLeuProAspLeuProGluGlnPheSerProProAspVal--- 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 538 GGAATGCCAGTCAGATTAAGCTTAAGGAGCTCCTTGAAGCAGAAAGATT 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 124 AlaProProIleLeuValLysLeu-----ValGluAlaIleGluArgThr 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 598 CCTCAATACACAGAGAGCTGATATCACTCTGTAGAAACCAAAATACTCAAGAG 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 139 GlyLeuAspSerTyrArgProGluProProAlaValArgThrAspPheSerLeuSerAsp 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 658 ATTCCA-----CTTGCATT 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 159 ValGluGlnTyrAspAlaAlaAlaLeuSerAspGlyValLysGlyPheLeuLeuAlaLeu 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 673 AGCAGTTCTTCATTCACGACAAAC----- 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 179 ProAlaProLeuValThrProGluAlaAlaAlaGluAlaHisArgAlaLeuArgGluAla 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 697 -----AACCAAGCTGTGCAAAAC 714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 199 AlaGlyValGlyProAlaLeuGluProProThrLeuProLeuHisAlaLeuThrLeu 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 715 AGA-----GATCATAGAGAGGAGCATGACCCCTGTTCTCCCTAG----- 753
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 219 ArgPheLeuLeuGlnHisLeuGlyArgValAlaGlyArgAlaProAlaProGlyProAla 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 754 -----AGATGCCAGCCTCCAGCAGCAG 774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 239 ValArgAlaLeuGlyAlaThrPheGlyProLeuLeuLeuArgAlaProProPhePro 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 775 TGCAGCCCT----- 783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259 ProProProGlyGlyAlaProAspGlyThrGluProThrProAspPheProAlaLeuLeu 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 784 -----CACGAATATACTGCTATATAATACACA 813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279 ValGluLysLeuLeuGlnGlnHisLeuGluGlnGlnValAlaLeuProAlaLeu--- 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 814 AGCTGGAGACACCTTTCCCAAAAGTCT----- 843
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

DB 298 -----ProProLysProProLysThrLysProAlaProThrGlyLeuAlaAsnGly 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 844 ---GATGAAGAAGATCTCCAGACCAATGATGTGATATTGGAGATACAGCCGCCAGCA 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 315 GlySerProProSerLeuAsnAlaGluTyrTyrGlyAspIleSerArgGluGln 334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 901 GTGAGAGAGCATTCATGAGAGAGCAAGCAAGATGATGTTCTGTGTCGAGATTGTCC 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 335 ValAsnGluLys---LeuArgAspThrProAspGlyThrPheLeuValArgAspAlaSer 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 961 ACAAAATCCAGAGAGAGACCCCTATGTTTGGCTGTGTTTAAAGACAAGATTCACAAAT 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 SerLysIleGlnGlyLys---TyrThrLeuThrLeu-----ArgLysGlyGlyAsn 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1021 GTTAAAT---ATCCGCTCTCTCGAGAGAGATTCAGCACTTTGCCCTGGGACAGACTCACA 1077
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 370 AsnLysLeuIleLysValPheHisArgAspGlyHisTyrGlyPheSerGluProLeu--- 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1078 GGAGATGAGAGATTGATTGATTCAGTGAAGACATCATGACACTACAGAAAT 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 -----ThrPheCysSerValValAspLeuIleThrHisTyrArgHis 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

RESULT 14

H59435
 phosphoinositide-3-kinase regulatory beta chain [imported] - human
 C:Species: Homo sapiens (man)
 C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: H59435; A59436
 R:Volinina S; Patrachini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L
 Oncogene 7, 789-793, 1992
 A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinositol

A:Reference number: H59435
 A:Accession: H59435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <VOL>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1
 R:Janssen, J W.; Schleithoff, L.; Barttram, C.R.; Schulz, A.S.
 Oncogene 16, 1767-1772, 1998
 A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta sub

A:Reference number: A59436; MUID:98241181; PMID:9582025
 A:Accession: A59436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <JAN>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1

Alignment Scores:

Pred. No.:	0.0413	Length:	728
Score:	126.50 <td>Matches:</td> <td>75 </td>	Matches:	75
Percent Similarity:	34.34% <td>Conservative:</td> <td>27 </td>	Conservative:	27
Best Local Similarity:	25.25% <td>Mismatches:</td> <td>108 </td>	Mismatches:	108
Query Match:	6.26% <td>Indels:</td> <td>88 </td>	Indels:	88
DB:	2 <td>Gaps:</td> <td>15 </td>	Gaps:	15

US-09-856-061-3 (1-1129) x H59435 (1-728)

```

QY 287 CCAGAGCTCTA---TCTCCATTGGACAGCCGACCTGGAACACACAGACGAGTTGGAAA 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 ProAlaProLeuValThrProGluAlaSerAlaGluAla-----ArgArgAlaLeuArg 199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 344 GAGTGGACAACCACTTTCAGGAGAGTCAGAAAGCCAAACATTAAAGAGATGATCCG 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 200 GluAlaAlaGlyProValGlyProAlaLeuGluProProThrLeu---ProLeuHisArg 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 404 TAAAGAAACAAGATCTTTCACACCTCTCGGCT----- 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 219 AlaLeuThrLeuArgPheLeuLeuGlnHisLeuGlyArgValAlaArgArgAlaProAla 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 442 -----CTCATTAACACTTCGAGAGAGATCAACACCTTG-----CCGCTT 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 239 LeuGlyProAlaValArgAlaLeuGlyAlaThrPheGlyProLeuLeuLeuArgAlaPro 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```



```

QY 481 GAGCCGAGACGACGACCTTTATCTCAGAGACACACCTTTCCAGAGATCCAGGA 540
    ||| ||||| |||||
Db 259 ProProSerSerProPro-----GlyGly 268
QY 541 ATGCCAGTCAGATTAAGCTTAAGGACTTAAGTAGCTCTTGAAGAGAAAAATTCT 600
    |||
Db 269 AlaProAsp----- 271
QY 601 CATAGCAGAGAGACCTGATCACTATCTGTAGAAAACCAAAATACTCAAGATT 660
    : : : : : ||||| : : :
Db 272 GlySerGluProSerProAspPheProAlaLeuLeuValGluLysLeuGluGlu--- 290
QY 661 CCAGTTCGATTAAGCAGTTCTTCATTCACAGACAAACACAGTGTGCAAAACAGAGAT 720
    : : : : :
Db 290 ----- 290
QY 721 CATAGAGAGAGCAGACCCCTGTTCTTCAGAGATGCCAGCTCCAGCCAGCTGCAGC 780
    : : : : : |||
Db 291 HisLeuGluGluGluValAlaProProAlaLeuProProLysProProLysAlaLys 310
QY 781 CCTCAGCAAAATATACGCTTATTAATACACAGCTGGAGACCACTTTCCCAAAAG 840
    ||| : : : : : |||||
Db 311 ProAlaProThrValLeuAla-----AsnGlyGlySerProPro----- 323
QY 841 TCTGATAGAAAGATGTCACACACATGATGATGATGAGATACAGCCGCGCAGCA 900
    : : : : : ||||| : : : : :
Db 324 -----SerLeuGlnAspAlaGluTrpTyrTrpGlyAspLysLeuGluGlu 339
QY 901 GTGCAAGAGCATTCATGAGAGAGACACAGATGCTGTTCTGTCGCGAGATTGTTCC 960
    ||| ||| : : : : : ||||| |||||
Db 340 ValAsnGluLys---LeuAlaArgAspThrProAspGlyThrPheLeuValAlaArgAspAlaSer 358
QY 961 ACAAAATCCAGAGAGACCCCTATGTTTGGCTGTGTTTATGAAACAAAGCTACAT 1020
    : : : : : ||| : : : ||| |||
Db 359 SerLysIleGluGluGlu---TyrThrLeuThrLeu-----ArgLysGlyGlyAsn 374
QY 1021 GTAAA---ATCCGCTTCCTGGAGAGATCAGCAGTTTGCCCTGGGACAGACTCAGA 1077
    ||| : : : : : ||||| : : :
Db 375 AsnLysLeuIleLysValAlaPheHisArgAspGlyHisTyrGlyPheSerGluProLeu--- 393
QY 1078 GGAGATGAGAGATTGATTCAGTAGAAGACATCTCCAGACCTACAGAAAT 1128
    ||| ||||| ||||| : : : : :
Db 394 -----ThrPheCysSerValValAspLeuIleAsnHisTyrArgHis 407

```

RESULT 15

S08636

Nck protein - human

N:Alternate names: src-related protein

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 05-Nov-1999

C:Accession: S08636

R:Lehmann, J.M.; Rietmuller, G.; Johnson, J.P.

Nucleic Acids Res. 18, 1048, 1990

A:Title: Nck, a melanoma cDNA encoding a cytoplasmic protein consisting of the src homo-

A:Reference number: S08636; MUID:90192089; PMID:2107526

A:Accession: S08636

A:Molecule type: mRNA

A:Residues: 1-377 <LEH>

A:Cross-references: EMBL:X17576; NID:q35014; PIDN:CAA35599.1; PID:q35015

C:Superfamily: SH3 homology; SH2 homology

F:9-56/Domain: SH3 homology <SH31>

F:113-160/Domain: SH3 homology <SH32>

F:197-247/Domain: SH2 homology <SH33>

F:282-371/Domain: SH2 homology <SH2>

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length: 377

Matches: 33

Conservative: 23

Mismatches: 52

Indels: 32

Gaps: 4

```

US-09-856-061-3 (1-1129) x S08636 (1-377)
QY 706 GTGCAAAACAGAGATCATAGAGAGCGATCGACCCCTGTTCTCTCAGAGATGCCAGCT 765
    : : : : : ||||| : : : : :
Db 250 MetGlnAsnAsnProLeuThrSerGlyLeuGluProSerProProGln----- 265
QY 766 CCAGCAGCTCAGACCCCTCAGAAATATATACTGGCCCTTAATATACACAGCTGGAGACA 825
    ||| ||| : : : : :
Db 266 -----CysAspTyrIleArgProSerLeuThrGlyLysPheAlaGly----- 279
QY 826 CTTTCCCAAAAGCTGTGATGAAAGAGATGCCAGCACAATGATGATGATGAGAA 885
    : : : : : ||| ||||| : : : : :
Db 280 -----AsnProTyrTyrGlyLys 286
QY 886 TACAGCCGCGCAGCAGTGAAGAGCATTCATGAAGAGAAACAAGATGATGTTCTTG 945
    : : : : : ||| ||| : : : : :
Db 287 ValThrArgHisGlnAlaGluMetAlaLeuAsnGluArgGlyHisGluLysAspPheLeu 306
QY 946 GTCCGAGATTGTTCCACAAATCCAGAGAGAGCCCTATGTTTGGCTGTGTTTATGAG 1005
    : : : : : ||| : : : : :
Db 307 IleArgAspSerGluSerSerProAsnAspPheSerValSerLeuLysAlaGluGlyLys 326
QY 1006 AACAAAGTCTCAATGTAAAAATCCGCTTCCTGGAGAGAGAAATCAGAGTTGCCCTGGG 1065
    ||||| : : : : : ||||| |||||
Db 327 AsnLysHisPheLysValGlnLeu-----LysGluThrValTyrCysIle--- 341
QY 1066 ACAGAGCTCAAGAGAGATGAGAAAGTTGATTCAGTAGAAGACATCATCGAAGCTACAG 1125
    ||| ||||| : : : : : ||||| |||||
Db 342 -----GlyGlnArgLysPheSerThrMetGluGluValGlnHisTyrLys 357

```

Search completed: April 21, 2003, 12:32:13

Job time : 48.2565 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:11 ; Search time 16.2418 seconds

(without alignments)
5766.221 Million cell updates/sec

Title: US-09-856-061-3
Perfect score: 2020
Sequence: 1 ttcacgaactcagctcgtcc.....catcgacactacaagaatt 1129

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame-n2p.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09856061/runat.21042003.113018.3735/app_query.fasta.1.3150
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09856061.ccgcn.1.1.65 @runat.21042003.113018.3735 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEROUT=120
-WARN_TIMEROUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332.5	16.5	533	1	LCP2_HUMAN
2	315.5	15.6	533	1	LCP2_MOUSE
3	154.5	7.6	724	1	P85A_BOVIN
4	150.5	7.5	724	1	P85A_HUMAN
5	149	7.4	724	1	P85A_MOUSE
6	145	7.2	724	1	P85A_RAT
7	143	7.1	659	1	BTK_MOUSE
8	140	6.9	559	1	BTK_MOUSE
9	140	6.9	659	1	BTK_HUMAN
10	136.5	6.8	645	1	CSW_MOUSE
11	136	6.7	561	1	3BP2_HUMAN
12	133.5	6.6	724	1	P85B_BOVIN
13	128	6.3	461	1	P55G_HUMAN
14	128	6.3	722	1	P85B_MOUSE
15	127	6.3	271	1	CRK_DROME
16	126.5	6.3	533	1	SOC6_MOUSE
17	126.5	6.3	728	1	P85B_HUMAN
18	126.5	6.3	1490	1	CRK7_HUMAN

19	126	6.2	461	1	P55G_BOVIN	046404 bos taurus
20	125	6.2	377	1	NCK1_HUMAN	P16333 homo sapien
21	125	6.2	585	1	PTNB_MOUSE	P35235 mus musculu
22	125	6.2	593	1	PTNB_HUMAN	006124 homo sapien
23	125	6.2	593	1	PTNB_RAT	P41499 ratius norv
24	124.5	6.2	217	1	GRAP_HUMAN	013588 homo sapien
25	123.5	6.1	722	1	P85B_RAT	063788 ratius norv
26	123.5	6.1	1196	1	ABLI_CAEEL	P03949 caenorhabd
27	123	6.1	461	1	P55G_MOUSE	064143 mus musculu
28	122	6.0	822	1	PER_HUMAN	P16591 homo sapien
29	120.5	6.0	893	1	YM92_CAEEL	P34531 caenorhabd
30	120	5.9	847	1	VAV2_HUMAN	09ukw4 homo sapien
31	120	5.9	939	1	ST20_YEAST	003497 saccharomyc
32	119	5.9	620	1	ITK_HUMAN	008881 homo sapien
33	119	5.9	1872	1	T2D1_HUMAN	P21675 homo sapien
34	116	5.7	517	1	FGR_MOUSE	P14234 mus musculu
35	115	5.7	675	1	BMX_HUMAN	P51813 homo sapien
36	115	5.7	1290	1	PIG1_HUMAN	P19174 homo sapien
37	114.5	5.7	625	1	ITK_MOUSE	003526 mus musculu
38	114.5	5.7	630	1	TEC_MOUSE	P24604 mus musculu
39	114	5.6	1128	1	BEM3_YEAST	P32873 saccharomyc
40	114	5.6	1190	1	ZO2_HUMAN	09udv2 homo sapien
41	114	5.6	2440	1	NCK1_HUMAN	075376 homo sapien
42	112.5	5.6	520	1	YEAE_SCHPO	014079 schizosacch
43	112.5	5.6	303	1	CRKL_HUMAN	P46109 homo sapien
44	112.5	5.6	878	1	VAV2_HUMAN	P52735 homo sapien
45	112.5	5.6	1942	1	Y054_HUMAN	P42694 homo sapien

ALIGNMENTS

RESULT 1
LCP2_HUMAN STANDARD; PRT; 533 AA.
ID LCP2_HUMAN
AC Q13054;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte
DE protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
GN LCP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Leukemia;
RX MEDLINE=95221345; PubMed=7706237;
RA Jackman J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A.,
RT Koretzky G.A., Findell P.R.;
RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein
RT associated with Grb2 in T cells.";
RL J. Biol. Chem. 270:7029-7032(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND Fyb.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND
CC PERIPHERAL BLOOD LEUCOCYTES. HIGHLY EXPRESSED ALSO IN T CELL AND
CC MONOCYTIC CELL LINES. EXPRESSED AT LOWER LEVEL IN B CELL LINES.
CC NOT DETECTED IN FIBROBLAST OR NEUROBLASTOMA CELL LINES.
CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION BY ZAP-70.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- DATABASE: NAME=PROW; NOTE=PROW 1.1:-5(2000);
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/1118450040_g.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: U20158; AAC50135.1; -
 DR EMBL: BC016618; AAH16618.1; -
 DR HSSP: P12931; 1SHD.
 DR GeneW: HGNC:6529; LCP2.
 DR MIM: 601603; -
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00454; SAM; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR SH2 domain; Phosphorylation.
 FT DOMAIN 133 136 POLY-GLU.
 FT DOMAIN 198 201 POLY-PRO.
 FT DOMAIN 422 530 SH2.
 SQ SEQUENCE 533 AA; 60188 MW; CSD22F31D36200C8 CRC64;

Alignment Scores:
 Pred. No.: 1,09e-17 Length: 533
 Score: 332.50 Matches: 113
 Percent Similarity: 36.24% Conservative: 49
 Best Local Similarity: 25.28% Mismatches: 128
 Query Match: 16.46% Indels: 157
 DB: Gaps: 16

US-09-856-061-3 (1-1129) x LCP2_HUMAN (1-533)

OY 58 GGCCGATCCAGAGATGATGAAGACCTCTTCTAGAGTGGGAAAGAACTTCTCTCAGTC 117
 DB 126 G1YASP1YRGLUSERPROASN-----GluGluGluGluAlaProVal 139
 OY 118 CTGATGAGCAAAAGGCCACATGATGATGATGATGATGATGATGATGATGATGATGAT 177
 DB 140 G1UASP-----AspAlaAsp1YRGLUPROPROPROSERASNspGlu 153
 OY 178 GAGACATGCGAGTGAATTAATTTTACCAGCGCGGCTATTAAGAAATGGAATGCA 237
 DB 154 G1UALALEuGlnAsn---Ser1LeuProAlaLysProPheProAsnSer----- 169
 OY 238 GATTAACACTATTTCAGAGTTTCATGACACTCCCCCTTCCCTTAGACACAGACCTCT 297
 DB 170 AsnSerMet1YR----- 173
 OY 238 ATCTCATTTGGACAGCGACCTGGAACACACAGAGAGTTTGAAAGAGTGACAAACC 357
 DB 174 -----1LeaspArgPro 177
 OY 358 ATTTCAGGAGCTCAGAACGCCAAACATTAAGAGATGATCGCTAAGAAAGAACAG 417
 DB 178 ProSerGly1YRThrProGlnGlnPro----- 186
 OY 418 ATTCTTTACCACTCCCTCGGCTCTCATACACTCCG-----AAGAAG 462
 DB 187 ---ProValProProGlnArgProMetAlaAlaLeuProProProProAlaGlyArgAsn 205
 OY 463 TACCAACCTTGCCCTC-----GAGCCGAGAGAGCAGAG----- 498
 DB 206 H1SSERProLeuProProProGlnThrAsnH1SGluGluProSerArgSerArgAsnH1S 225
 OY 498 ----- 498
 DB 226 LysThrAlaLysLeuProAlaProSer1LeaAspArgSerThrLysProProLeuAspArg 245
 OY 499 -----CCACCTTTATCTCAG 513
 DB 246 Ser1LeuAlaProPheAspArgGluProPheThrLeuGlyLysLysLysProProPheSerAsp 265

OY 514 AGACACACCTTTCCA-----GAAGTCCAG 537
 DB 266 LysProSer1LeuProAlaGlyArgSer1LeuGlyGlnH1SLeuProLys1LeuGlnLysPro 285
 OY 538 GGAATGCCCACTGATGAATGAAGGACTTAAGGAGTCTTAAGTGAAGTCTTGAAGCAGAAAGT 597
 DB 286 ProLeuProProThrThrGlnArgH1SGluArgSerSerProLeuProGlyLysLysPro 305
 OY 598 CCT-----CAT 603
 DB 306 ProValProLysH1SGlyTrpGlyProAspArgGluAsnAspGluAspValH1S 325
 OY 604 AACCAAGAGAACCTGATCACTCATCTGTAGAAACCAAAATACT----- 651
 DB 326 GlnArgProLeuProGlnProAlaLeuLeuProMetSerSerAsnThrPheProSerArg 345
 OY 652 -----CAAGAGATTCCACTTCCCATTAAGCAGTTCTTCAATTCAG 690
 DB 346 SerThrLysProSerProMetAsnProLeuProSerSerH1SMeLeuProGlyAlaPheSer 365
 OY 691 ACAAGCAACCAAGTGTGCAAAACAGAGATCATTAAGAGGAGCATGACGCCCTGTCTCT 750
 DB 366 GluSerAsnSerSerPhe-----ProGlnSerAlaSerLeuProProTyrrPheSer 382
 OY 751 CAG-----AGATGCCAGCCTCCAGCCAGCTGACAGCCCTCAGCAAAATATATCTGCCCTAT 804
 DB 383 GlnGlyProSerAsnArgProProLeuArgAlaGluGlyArgAsnPheProLeuProLeu 402
 OY 805 AAATACACAAAGTGGAGACACACCTTCCCAAAAGTGTGATGAAGAGATGTCACAGAC 864
 DB 403 Pro---AsnLysProArgProProSerPro-----AlaGluGluGluAsnSerLeuAsn 419
 OY 865 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 924
 DB 420 GlnGluTrp1YRValSer1YR1LeuThrArgProGluAlaGluAlaLeuArgLys1Le 439
 OY 925 AACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 984
 DB 440 AsnGlnAspGlyThrPheLeuValArgAspSerSerLysThrThrThrAsnProTyrr 459
 OY 985 GTTTGGCTGTGTTTATGAGAACAAAGTCTAATGTAAATAATCCCTCTCTGAGAGG 1044
 DB 460 ValLeuMetValLeuLysAspLysVal1YRAsn1LeuGln1LeuArg1Gln1LysGlu 479
 OY 1045 AATCAGCACTTGGCCCTGGGAGCAGCTCAGAGAGATGAGAAATGATGATGATGATGATGAT 1104
 DB 480 SerGlnVal1YRLeuLysGlyThrGlyLeuArgGlyLysGluAspPheLeuSerValSer 499
 OY 1105 GACATCATGACACTACAG 1125
 DB 500 Asp1Leu1LeuAsp1YR1PheArg 506

RESULT 2
 LCP2_MOUSE STANDARD: PRT: 533 AA.
 AC 060787;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte
 DE protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
 GN LCP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC TISSUE=T-cell lymphoma;
 RX MEDLINE=95221345; PubMed=7706237;
 RA Jackson J.K., Motlo D.G., Sun Q., Tanemoto M., Turk C.W., Pelicz G.A.,
 RA Koretzky G.A., Finkel P.R.;

RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Larrea F., Panayotou G.,
 RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
 RA Courtneidge S.A., Parker P.J., Waterfield M.D.;
 RT "Characterization of two 85 kd proteins that associate with receptor
 RT tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase.";
 RL Cell 65:91-104(1991).
 [12]
 RP CIRCULAR DICHROISM AND FLUORESCENCE SPECTROSCOPY.
 RX MEDLINE=93049176; PubMed=1330535;
 RA Panayotou G., Bax B., Gout I., Federwisch M., Wrzobowski B., Dhand R.,
 RA Fry M.J., Blundell T.L., Moliner A., Waterfield M.D.;
 RT "Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
 RT domain with a PDGF receptor phosphorylation site: structural features
 RL and analysis of conformational changes.";
 EMBO J. 11:4261-4272(1992).
 [13]
 RP STRUCTURE BY NMR OF 1-84.
 RX MEDLINE=93272320; PubMed=7684655;
 RA Bocker G.W., Gout I., Downing A.K., Driscoll P.C., Boyd J.,
 RA Waterfield M.D., Campbell I.D.;
 RT "Solution structure and ligand-binding site of the SH3 domain of the
 RT p85 alpha subunit of phosphatidylinositol 3-kinase.";
 RL Cell 73:813-822(1993).
 [14]
 RP STRUCTURE BY NMR OF 314-431.
 RX MEDLINE=92357146; PubMed=1323062;
 RA Bocker G.W., Breeze A.L., Downing A.K., Panayotou G., Gout I.,
 RA Waterfield M.D., Campbell I.D.;
 RT "Structure of an SH2 domain of the p85 alpha subunit of
 RT phosphatidylinositol-3-OH kinase.";
 RL Nature 358:684-687(1992).
 [15]
 RP STRUCTURE BY NMR OF 321-434.
 RX MEDLINE=97110350; PubMed=8952511;
 RA Guenther U.L., Liu Y., Sanford D., Bachovchin W.W., Schaffhausen B.;
 RT "NMR analysis of interactions of a phosphatidylinositol 3'-kinase SH2
 RT domain with phosphotyrosine peptides reveals interdependence of major
 RT binding sites.";
 RL Biochemistry 35:15570-15581(1996).
 [16]
 RP STRUCTURE BY NMR OF 614-724.
 RX MEDLINE=96173872; PubMed=9512716;
 RA Siegal G., Davis B., Kristensen S.M., Sankar A., Linacre J.,
 RA Stein R.C., Panayotou G., Waterfield M.D., Driscoll P.C.;
 RT "Solution structure of the C-terminal SH2 domain of the p85 alpha
 RT regulatory subunit of phosphoinositide 3-kinase.";
 RL J. Mol. Biol. 276:461-478(1998).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNTS.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M61745; AAA79511.1; -;
 DR PIR: A38749; A38749.
 DR PDB: 2PNA; 31-JAN-94.
 DR PDB: 2PNB; 31-JAN-94.
 DR PDB: 2PNJ; 31-OCT-93.
 DR PDB: 1PNJ; 31-OCT-93.
 DR PDB: 1BFI; 25-FEB-98.

DR PDB: 1BFI; 25-FEB-98.
 DR InterPro: IPR001720; PI3kinase_p85.
 DR InterPro: IPR000198; RHOGAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RHOGAP; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RHOGAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS0001; SH2; 2.
 DR PROSITE: PS0002; SH3; 1.
 KW SH3 domain; SH2 domain; Repeat: 3d-structure.
 FT DOMAIN 3
 FT DOMAIN 129 273
 FT DOMAIN 333 428
 FT DOMAIN 624 718
 FT HELIX 340 347
 FT TURN 348 348
 FT STRAND 354 359
 FT STRAND 368 373
 FT STRAND 378 382
 FT STRAND 384 384
 FT STRAND 391 391
 FT STRAND 391 391
 FT HELIX 401 409
 FT STRAND 413 413
 FT TURN 418 419
 FT STRAND 422 422
 FT STRAND 427 427
 SQ SEQUENCE 724 AA; 83497 MW; EBD6E6754BBF7321 CRC64;
 Alignment Scores:
 Pred. No.: 0.000324 Length: 724
 Score: 154.50 Matches: 81
 Percent Similarity: 37.21% Conservative: 47
 Best Local Similarity: 23.5% Mismatches: 102
 Query Match: 7.65% Indels: 114
 DB: Caps: 14
 US-09-856-061-3 (1-1129) x P85A_BOVIN (1-724)
 QY 406 AGAAGAACAAAGATCTCTTACCACTCTCGGCTCTCATTAACACTTCCGACAGATAC 465
 Db :::: 111 111 111 111 111
 Db 80 LysLysLysSerProProThrProLysProArGProProArGProLeuProValAlaPro 99
 QY 466 CAACCCCTGGCCCGGACGCGGAGAGACGAGG-----CCACCTTATCT 510
 Db :::: 111 :::: :::: :::: ::::
 Db 100 GlyProSerLysThrGlnAlaAspSerGlnGlnAlaSerThrLeuProAspLeuAla 119
 QY 511 CAGAGACACACCTTCCGAGAGTC----- 534
 Db :::: :::: :::: :::: ::::
 Db 120 GluGlnPheAlaProProAspValAlaProProLeuLeuIleLysLeuValGlnAlaIle 139
 QY 535 -----CAGGGAAGGCCCAAGTCAGATTAAGCTTAAG----- 564
 Db :::: :::: :::: :::: ::::
 Db 140 GluLysLysGlyLeuGluCysSerThrLeuTyrGlnSerSerSerAsnProAla 159
 QY 565 GACCTTAATGAGTCCTTGAGCAGCAAAAGTCTCT----- 600
 Db :::: :::: :::: :::: ::::
 Db 160 GluLeuArgGlnLeuLeuAspCysAspThrAlaSerLeuAspLeuGluMetPheAspVal 179
 QY 601 CATAAACGAGAGCAAGCCTGAATCAACATCTGTTAGAAAACAAATATCTCAAGACATT 660
 Db :::: :::: :::: :::: ::::
 Db 180 HisValLeuAlaAspAlaPheLysArgTyrLeuLeuAspLeuProAsn---ProValIle 198
 QY 661 CCACCTGCGATTAACAGCTCT----- 681
 Db :::: :::: :::: :::: ::::
 Db 199 ProValAlaValSerSerGlnLeuIleSerLeuAlaProGluValGlnSerSerGlnGlu 218

QY 681 ----- 681
 Db 219 TyrIleGlnLeuLeuLysLysLeuIleArgSerProSerIleProHisIleTyrTrpLeu 238
 QY 682 ----- TCATTCACGACAGCAACACACACT 705
 Db 239 ThrLeuGlnTyrLeuLeuLysHisPhePheLysLeuSerGlnThrSerSerLysAsnLeu 258
 QY 706 GTGCAAAACAGATCATGATGAGGACATGACGACCTGTTCTCCACAGATGCCACCT 765
 Db 259 LeuAsnAlaArg-ValLeuSerGlnLeuPheSerProLeuLeuPheArg-PheProAla 278
 QY 766 CCAGCCAGCTGCAGCCCTCAGCAAAATATACCTGCCCTATTAATAC-----ACAAAGC 816
 Db 278 laSer-SerGlnAsnThrGlnHisLeuIleLysIleIleGlnIleLeuIleSerThrGlu 297
 QY 817 TGGAGA-----CCACCTTTCACAAAG----- 840
 Db 298 TrpAsnGluArgGlnProAlaProAlaLeuProProLysProProLysProThrThrVal 317
 QY 841 -----TCGTATGAAAAGATGTCAGACAAATGAATGTFACATTTGGAGAA 885
 Db 318 AlaAsnAsnGlyMetAsnAsnAsnMetSerLeuGlnAspAlaGlnTrpTyrTrpGlyAsp 337
 QY 886 TACAGCCCGCAGGAGTGAGAGGACATTCATGAAGAGACAGAGATGATGTTCTTG 945
 Db 338 IleSerArgGlnGluValAsnGlnLys--LeuArgAspThrAlaAspLysThrPheLeu 356
 QY 946 GTCCAGATGTTCACCAAAATCCAGAGAGACCCATGTTGGCTGTGTTTATGAG 1005
 Db 357 ValArgAspAlaSerThrLysMetLysGlyAsp--TrpThrLeuThrLeu----- 372
 QY 1006 AACAAAGTCTACATGTATAA--ATCCGCTTCCTGAGAGAGATGACAGATTTGCCCTG 1062
 Db 373 ArgLysGlyGlyAsnAsnLysLeuIleLysIlePheHisArgAspLysTyrGlyPhe 392
 QY 1063 GGGACAGACCTCAGAGAGATGAGAATTGATTGATGAGAACATCATCGAACACTAC 1122
 Db 393 SerAspProLeu-----ThrpAsnSerValValGlnLeuIleAsnHisTyr 408
 QY 1123 AAGANT 1128
 Db 409 ArgAsn 410
 RESULT 4
 P85A_HUMAN STANDARD; PRT; 724 AA.
 AC P27986;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
 DE p85-alpha subunit) (PtdIns-3-kinase p85-alpha) (PI3K).
 GN PI3K1 OR GRB1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Primates; Carnivora; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91191565; PubMed-1849461;
 RX Skolnik E.Y., Margolis B., Mohammadi M., Lowenstein E., Fischer R.,
 RA Drepps A., Ullrich A., Schlessinger J.;
 RT "Cloning of p13 kinase-associated p85 utilizing a novel method for
 RT expression/cloning of target proteins for receptor tyrosine
 RT kinases.";
 RL Cell 65:83-90(1991).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 1-85.
 RX MEDLINE-96196433; PubMed-8648629;
 RA Liang J., Chen J.K., Schreiber S.L., Clardy J.;
 RT "Crystal structure of PI3K SH3 domain at 2.0-A resolution.";

RL J. Mol. Biol. 257:632-643(1996).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 324-434.
 RX MEDLINE-96185448; PubMed-8599763;
 RA Nolte R.T., Eck M.J., Schlessinger J., Shoelson S.E., Harrison S.C.;
 RT "Crystal structure of the PI 3-kinase p85 amino-terminal SH2 domain
 RT and its phosphopeptide complexes.";
 RL Nat. Struct. Biol. 3:364-373(1996).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 115-298.
 RX MEDLINE-97121392; PubMed-8962058;
 RA Musacchio A., Cantley L.C., Harrison S.C.;
 RT "Crystal structure of the breakpoint cluster region-homology domain
 RT from phosphoinositide 3-kinase p85 alpha subunit.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:14373-14378(1996).
 RN [5]
 RP STRUCTURE BY NMR OF 1-79.
 RX MEDLINE-93208889; PubMed-7681364;
 RA Koyama S., Yu H., Dalgarno D.C., Shin T.B., Zydowsky L.D.,
 RA Schreiber S.L.;
 RT "Structure of the PI3K SH3 domain and analysis of the SH3 family.";
 RL Cell 72:945-952(1993).
 RN [6]
 RP STRUCTURE BY NMR OF 91-104.
 RX MEDLINE-97121261; PubMed-8961927;
 RA Renzoni D.A., Pugh D.J., Siliqardi G., Das P., Morton C.J., Rossi C.,
 RA Waterfield M.D., Campbell I.D., Ladbury J.E.;
 RT "Structural and thermodynamic characterization of the interaction of
 RT the SH3 domain from Fyn with the proline-rich binding site on the p85
 RT subunit of PI3-kinase.";
 RL Biochemistry 35:15646-15653(1996).
 RN [7]
 RP STRUCTURE BY NMR OF 617-724.
 RX MEDLINE-96312955; PubMed-8670861;
 RA Breeze A.L., Kara B.V., Barratt D.G., Anderson M., Smith J.C.,
 RA Luke R.W., Best J.R., Cartledge S.A.;
 RT "Structure of a specific peptide complex of the carboxy-terminal SH2
 RT domain from the p85 alpha subunit of phosphatidylinositol 3-kinase.";
 RL EMBO J. 15:3579-3589(1996).
 RN [8]
 RP VARIANT ILE-326.
 RX MEDLINE-97184306; PubMed-9032108;
 RA Hansen T., Andersen C.B., Ewald S.M., Urhammer S.A., Clausen J.O.,
 RA Vestergaard H., Owens D., Hansen L., Pedersen O.;
 RT "Identification of a common amino acid polymorphism in the p85alpha
 RT regulatory subunit of phosphatidylinositol 3-kinase: effects on
 RT glucose disappearance constant, glucose effectiveness, and the
 RT insulin sensitivity index.";
 RL Diabetes 46:494-501(1997).
 RN [9]
 RP VARIANT INSULIN RESISTANCE GLN-409, AND VARIANT ILE-326.
 RX MEDLINE-20330645; PubMed-10768093;
 RA Baynes K.C.R., Beeton C.A., Panayotou G., Stein R., Soos M.,
 RA Hansen T., Simpson H., O'Rahilly S., Shepherd P.R., Whitehead J.P.;
 RT "Natural variants of human p85 alpha phosphoinositide 3-kinase in
 RT severe insulin resistance: a novel variant with impaired
 RT insulin-stimulated lipid kinase activity.";
 RL Diabetologia 43:321-331(2000).
 CC -I- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -I- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNIT.
 CC -I- DISEASE: DEFECTS IN PI3K1 ARE A CAUSE OF SEVERE INSULIN
 CC RESISTANCE.
 CC -I- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC DATABASE: NAME=PI3K; NOTE=PI3K 1.6-12(2000);
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1773542685_g.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).

DR EMBL; M61906; ?; NOT_ANNOTATED_CDS.
 DR PIR; A38748; A38748.
 DR PDB; 1PKS; 31-MAY-94.
 DR PDB; 1PKT; 31-MAY-94.
 DR PDB; 1PHT; 07-DEC-95.
 DR PDB; 1PBM; 12-MAR-97.
 DR PDB; 1PIC; 17-SEP-97.
 DR PDB; 1A0N; 25-FEB-98.
 DR PDB; 1AZG; 25-FEB-98.
 DR Genew; HGNC:8979; PIK3R1.
 DR MIM; 171833; ?
 DR InterPro; IPR001720; PI3Kinase_P85.
 DR InterPro; IPR000198; RhogAP.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00017; SH2; 2.
 DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00620; RhogAP; 1.
 DR PRINTS; PR00678; PI3KINASEP85.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR ProDom; PD000066; SH3; 1.
 DR ProDom; PD000093; SH2; 2.
 DR SMART; SM00324; RhogAP; 1.
 DR SMART; SM00352; SH2; 2.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS50001; SH2; 2.
 DR PROSITE; PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat: 3D-structure; Polymorphism;
 KW Disease mutation.
 FT DOMAIN 3 79 SH3.
 FT DOMAIN 333 428 SH2 1.
 FT DOMAIN 624 718 SH2 2.
 FT VARIANT 326 326 M -> I.
 FT VARIANT 409 409 R -> Q (IN SEVERE INSULIN RESISTANCE;
 FT REDUCTION OF INSULIN-STIMULATED
 FT ACTIVITY).
 FT FTID-VAR_010024.
 FT SEQUENCE 724 AA; 83597 MW; 5D9493C0431C760D CRC64;
 Alignment Scores:
 Pred. No.: 0.00065 Length: 724
 Score: 150.50 Matches: 82
 Percent Similarity: 37.04% Conservative: 48
 Best Local Similarity: 23.36% Mismatches: 93
 Query Match: 7.45% Indels: 128
 DB: 1 Gaps: 17

US-09-856-061-3 (1-1129) x P85A_HUMAN (1-724)

QY 406 AGAAGAACAAGATTCTTACACCTCTGCGCTCTCATACACTTCGAGAGAAGTAC 465
 : : : : :
 DB 80 Lysylsilesrprothrprolyrprolyrpro-----ProArg----- 93

QY 466 CAACTTCCTCCCTGAGCCGAGAGAGAGAG----- 498
 : : : : :
 DB 94 ---ProleuprovalalalprogllyserSerlystrglualalaspvalalgluglinala 112

QY 499 -----CCACTTTATCTCAGACACACCTTTCAGAGATC----- 534
 : : : : :
 DB 113 LeuthleuproaspLeuAlaIleglulnhealaproProaspIleAlaproProleu 132

QY 535 -----CAGGGAATG----- 543
 : : : : :
 DB 133 IlelyleuValalglualalaleglulnlyleuGlucylSerThrleuTyrrArgThr 152

QY 544 CCCAGTCATAGCTTAAAGACTTAGTGAGGCTCTTGAAGCAAGAAAGTCCAT 603
 : : : : :
 DB 153 GlnserSerSerAlaIleAlaIleuArgIleuLeuAspCysAsp---ThrProSer 171

QY 604 AACCAAGAGACCTGTAATCACTCATCTGTTAGAAACCA----- 645
 : : : : :
 DB 172 ValAspleuGlumetIleAspValAlaIleValLeuAlaAspAlaPheLysArgTyrrLeu 191

QY 646 -----AATACTCAAGATTCCTCCCTTTCAGCATCTCT----- 681
 : : : : :
 DB 192 AspleuProaspProvalIleleProAlaAlaValTyrrSerGlumetIleSerLeuAlaPro 211

QY 681 ----- 681
 : : : : :
 DB 212 GluValGlnSerSerGlulnTyrrIleGlnLeuLeuLysLysLeuIleArgSerProSer 231

QY 682 -----TCA 684
 : : : : :
 DB 232 IleProHISGlnTyrrTrpleuThrleuGlnTyrrleuLeuLysHISPhePheLysLeuSer 251

QY 685 TTCACGACACCAACACAGCTGTGCAAAACAGATCTCTACAGAGAGCATGCAGCCCTGT 744
 : : : : :
 DB 252 GlnThrSerSerLysAsnLeuLeuAsnAlaArg-ValLeuSerGlulnIlePheSerProme 271

QY 745 TCTCCTCAGA-----GATGCCAGCTCCAGCCAGCTGCAGCCCTCAGCAAA---ATATA 795
 : : : : :
 DB 271 lleuPheArgPheSerAlaAlaIleSerAspAsnThrGlusnleuIleLysValIleG1 291

QY 796 CTGCCCTATAAATACACAGCTGAGAA-----CCACCTTTC 831
 : : : : :
 DB 291 uileleuIleSerThrGlu---TrpAsnGluArgIleProAlaProAlaLeuProProLys 310

QY 832 CCCAAAGG-----TCTGATACAAAGATGCTCAGCAGC 864
 : : : : :
 DB 311 ProProLysProThrThrValAlaAsnAsnGlyMetAsnAsnMetSerLeuGlnAsn 330

QY 865 AATGATGTCATGTCAGTAATACAGCCGCGAGCAGTGAAGAGCATTCATAGAGAG 924
 : : : : :
 DB 331 AlaGlnTrpTyrrTrpGlyAspIleSerArgGluGluValAsnGluLys---LeuArgAsp 349

QY 925 AACCAAGATGTCATGTCCTGTCGAGATTCCTCCACAAATCCAAAGAGAGCCCTAT 984
 : : : : :
 DB 350 ThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHISglYAsp---Tyrr 368

QY 985 GTTTCGCTGTTTATGAGACAAAGTCTACATGTAA---ATCCGCTTCCTGAG 1041
 : : : : :
 DB 369 ThrleuThrleu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIlePheHIS 385

QY 1042 AGGATTCAGAGTTTCCCTGGGAGAGAGACTCAGAGAGAGATGAGAGTTTGTTCAGTA 1101
 : : : : :
 DB 386 ArgAspGlyLysTyrrGlyPheSerAspProleu-----ThrPheSerSerVal 401

QY 1102 GAAGACATCATCAGAACTACACAT 1128
 : : : : :
 DB 402 ValGlnleuIleAsnHISTyrrArgAsn 410

RESULT 5
 P85A_MOUSE
 ID P85A_MOUSE STANDARD: PRT: 724 AA.
 AC P26450;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
 DE p85-alpha subunit) (Pcdins-3-kinase p85-alpha) (PI3K).
 GN PI3K1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

CC STRAIN-BALB/C;
 RX MEDLINE-91191564; PubMed-1849460;
 RA Escobedo J.A.; Navankasatutas S.; Kavanaugh W.M., Milfay D.,
 RT Fried V.A.; Williams L.T.;
 RT "cDNA cloning of a novel 85 kd protein that has SH2 domains and
 RT regulates binding of p13-kinase to the PDGF beta-receptor.";
 RL Cell 65:75-82(1991).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC -----
 DR EMBL: M60651; AAA39886.1; -
 DR HSSP: P23727; 1BFI.
 DR MGD: MG1:97583; PI3K1.
 DR InterPro: IPR001720; PI3kinase_P85.
 DR InterPro: IPR000198; RhogAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PI3KINSEP85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 3 79 SH3.
 FT DOMAIN 333 428 SH2 1.
 FT DOMAIN 624 718 SH2 2.
 SQ SEQUENCE 724 AA; 83414 MW; DEAA6EE297CF07A CRC64;

Alignment Scores:
 Pred. No.: 0 000845 Length: 724
 Score: 149.00 Matches: 76
 Percent Similarity: 32.76% Conservative: 39
 Best Local Similarity: 21.65% Mismatches: 96
 Query Match: 7.38% Indels: 140
 DB: 1 Gaps: 17

US-09-856-061-3 (1-1129) x p85A_MOUSE (1-724)

QY 421 CCTTACACACCTCCGCTCTCATACACTTCCAGAGATGACCAACCTTGCCCTT 480
 DB 85 ProthProlysrProargPro-----Proarg-----ProleuproVal 97
 QY 481 GACCCGAGACGACGAG-----CCACCT 504
 DB 98 AlAPrCIGlSerSerlysrThrglualAspThrglInglInalaleuProleuproasp 117
 QY 505 TTATCTCAGAGACACACCTTCCAGAGATGCCAGGATGCCAGTACATACCTTAAGG 564
 DB 118 LeuAlaIeGlInHeAlaProProAspVal---AlaProProleuLeuIlelyLeuLeu 136

QY 565 GACTTATGAGAGTC-----CTTGAGACGAGAAAAGTTCCTCATATACAGAGAAACCT 618
 DB 137 GluAlaIleGlInuLySrlysglyLeuGlyCysSerThrLeuThrArgThrInSerSer 156
 QY 619 GAATCAACATCTGTGTAGAA-----AACCAATACTCTCAAGAG 657
 DB 157 AsnProAlaGlInuLeuArgInuLeuAspCysAspAlaAlaSerValAspLeuGluMet 176
 QY 658 ATTCACCTTGCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 717
 DB 177 IleAspValHisValLeuAlaAspAlaPheIlysrTyrlLeuAlaAspLeuProAsn 195
 QY 718 GATCATAGAGAGGATGACGACCCCTGTCTCTCT----- 750
 DB 196 -----ProValIleProValAlaValTyrlAsnGluMetSer 208
 QY 751 -----CAGAGATGCGAGCTCCAGCAGCTGC----- 777
 DB 209 LeuAlaIleGlInuLeuInSerProGluAspCysIleGlInuLeuLySylLeuIleArg 228
 QY 778 -----AGCCCTCAGCAAAATATACGCTCATATAATAC----- 810
 DB 229 LeuProAsnIleProHisGlyCysTrpLeuThrLeuGlnTrpLeuLeuHisPhePhe 248
 QY 810 -----ACAAGTGS----- 819
 DB 249 LysLeuSerGlnAlaSerSerLysAsnLeuAlaValAlaArgValLeuSerGluIlePhe 268
 QY 810 ----- 810
 DB 269 SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGlnHisLeuIleLys 288
 QY 811 -----ACAAGTGS----- 819
 DB 289 AlaIleGlInuLeuIleSerThrGluTrpAsnGluArgGlnProAlaProAlaLeuPro 308
 QY 820 ---AGACCACTTTTCCC-----AAAAGTCTGTATGAAAGATGTC 858
 DB 309 ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328
 QY 859 CAGCACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 918
 DB 329 GlnAspAlaGluTrpTyrlTrpGlyAspIleSerArgGluInuValAsnGluLys---Leu 347
 QY 919 AAGGAGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 978
 DB 348 ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrIlysrMetHisGlyAsp 367
 QY 979 CCTATGTTTGGCTGTGTTTATGAGAACCAAGTCTACATGTAA---ATCCGCTTC 1035
 DB 368 ---TyrThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIle 383
 QY 1036 CTGAGAGACATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
 DB 384 PheHisArgAspLysIlysrTyrlGlyPheSerAspProLeu-----ThrPheAsn 399
 QY 1096 TCAGTAGAGACATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
 DB 400 SerValValGluLeuIleAsnHisTyrlArgAsn 410

RESULT 6
 P85A.RAT STANDARD: PRT; 724 AA.
 ID P85A.RAT
 AC 063787; 063790; P70544; 055085;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-Kinase
 DE p85-alpha subunit) (PtdIns-3-kinase p85-alpha) (PI3K).
 GN PI3K1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116.

[1]

SEQUENCE FROM N.A. (ISOFORMS P85-ALPHA AND P55-ALPHA).

STRAIN=Wistar; TISSUE=Brain;

MEDLINE=96214979; PubMed=8621382;

Inukai K., Anai M., Van Breda E., Hosaka T., Katagiri H., Funaki M.,

Fukushima Y., Ogihara T., Yazaki Y., Kituchi M., Oka Y., Asano T.;

"A novel 55-kDa regulatory subunit for phosphatidylinositol 3-kinase

structurally similar to p55PIK is generated by alternative splicing

of the p85alpha gene.";

J. Biol. Chem. 271:5317-5320(1996).

[2]

SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).

TISSUE=Liver;

MEDLINE=97218222; PubMed=9065454;

Inukai K., Funaki M., Ogihara T., Katagiri H., Kanda A., Anai M.,

Fukushima Y., Hosaka T., Suzuki M., Shln B., Takata K., Yazaki Y.,

Kituchi M., Oka Y., Asano T.;

"p55alpha gene generates three isoforms of regulatory subunit for

phosphatidylinositol 3-kinase (PI 3-kinase), p50alpha, p55alpha, and

p85alpha, with different PI 3-kinase activity elevating responses to

insulin".

J. Biol. Chem. 272:7873-7882(1997).

[3]

SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).

TISSUE=Liver;

MEDLINE=97079666; PubMed=8921377;

Fruman D.A., Cantley L.C., Carpenter C.L.;

"Structural organization and alternative splicing of the murine

phosphoinositide 3-kinase p85 alpha gene.";

Genomics 37:113-121(1996).

-1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE

KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING

THE ASSOCIATION OF THE PILO CATALYTIC UNIT TO THE PLASMA MEMBRANE.

NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE

AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES

-1- SUBUNIT: HETEROIMER OF A PILO (CATALYTIC) AND A P85 (REGULATORY)

SUBUNITS.

-1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: P85-ALPHA (SHOWN HERE), P55-

ALPHA AND P50-ALPHA; ARE PRODUCED BY ALTERNATIVE SPLICING.

-1- TISSUE SPECIFICITY: THE P85-ALPHA ISOFORM IS WIDELY EXPRESSED.

EXPRESSION OF THE P55-ALPHA ISOFORM IS HIGHEST IN BRAIN AND

SKELETAL MUSCLE. THE P50-ALPHA ISOFORM IS ABUNDANT IN LIVER WITH

LOWER LEVELS IN BRAIN AND MUSCLE.

-1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.

-1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.

-1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch).

EMBL; D64045; BAA18932.1; -

EMBL; D64048; BAA18933.1; -

EMBL; U50412; AAC52846.1; -

EMBL; D78486; BAA24426.1; -

HSSP; P23727; 1BFI.

InterPro: IPR001720; PI3KInase_P85.

InterPro: IPR000198; RhOGAP.

InterPro: IPR000980; SH2.

InterPro: IPR001452; SH3.

Pfam; PF00017; SH2; 2.

Pfam; PF00018; SH3; 1.

Pfam; PF00620; RhOGAP; 1.

PRINTS; PR00678; PI3KINASEP85.

PRINTS; PR00401; SH2DOMAIN.

ProDom; PD000066; SH3; 1.

ProDom; PD000093; SH2; 2.

SMART; SM00324; RhOGAP; 1.

DR	SMART	SM00252	SH2	2
DR <td>SMART</td> <td>SM00326</td> <td>SH3</td> <td>1</td>	SMART	SM00326	SH3	1
DR <td>PROSITE</td> <td>PS50001</td> <td>SH2</td> <td>2</td>	PROSITE	PS50001	SH2	2
DR <td>PROSITE</td> <td>PS50002</td> <td>SH3</td> <td>1</td>	PROSITE	PS50002	SH3	1
KM <td>SH3 domain</td> <td>SH2 domain</td> <td>Repeat</td> <td>Alternative splicing</td>	SH3 domain	SH2 domain	Repeat	Alternative splicing
FT <td>DOMAIN</td> <td>3</td> <td>79</td> <td>SH3</td>	DOMAIN	3	79	SH3
FT <td>DOMAIN</td> <td>333</td> <td>428</td> <td>SH2 1</td>	DOMAIN	333	428	SH2 1
FT <td>DOMAIN</td> <td>624</td> <td>718</td> <td>SH2 2</td>	DOMAIN	624	718	SH2 2
FT <td>VARSPLIC</td> <td>1</td> <td>270</td> <td>MISSING (IN ISOFORM P55-ALPHA).</td>	VARSPLIC	1	270	MISSING (IN ISOFORM P55-ALPHA).
FT <td>VARSPLIC</td> <td>271</td> <td>304</td> <td>VLFEPFAPSNTLHLIKAVELLISAEWSEKOPA -> MY</td>	VARSPLIC	271	304	VLFEPFAPSNTLHLIKAVELLISAEWSEKOPA -> MY
FT <td>VARSPLIC</td> <td>1</td> <td>300</td> <td>ITVMEDIEDCAKTDINCGDILMFYIEMDP (IN</td>	VARSPLIC	1	300	ITVMEDIEDCAKTDINCGDILMFYIEMDP (IN
FT <td>VARSPLIC</td> <td>301</td> <td>306</td> <td>ISOFORM P55-ALPHA).</td>	VARSPLIC	301	306	ISOFORM P55-ALPHA).
FT <td>VARSPLIC</td> <td>301</td> <td>306</td> <td>MISSING (IN ISOFORM P50-ALPHA).</td>	VARSPLIC	301	306	MISSING (IN ISOFORM P50-ALPHA).
FT <td>VARSPLIC</td> <td>301</td> <td>306</td> <td>KOPAPA -> MHNLOT (IN ISOFORM P50-ALPHA)</td>	VARSPLIC	301	306	KOPAPA -> MHNLOT (IN ISOFORM P50-ALPHA)
SO <td>SEQUENCE</td> <td>724 AA</td> <td>83531 MW</td> <td>95C65CF612873B84 CRC64</td>	SEQUENCE	724 AA	83531 MW	95C65CF612873B84 CRC64

Alignment Scores:	Pred. No.:	Score:	Length:	Matches:
Percent Similarity:	0.00169	145.00	724	93
Best Local Similarity:	35.43%	35.43%	Conservative:	48
Query Match:	23.37%	23.37%	Mismatch:	155
DB:	7.18%	7.18%	Indels:	102
	1	1	Gaps:	18

US-09-856-061-3 (1-1129) x P85A_RAT (1-724)	US-09-856-061-3 (1-1129) x P85A_RAT (1-724)
19	CGAAAAACAGGTCTAGCGCTCGCATCAATGATGGCCACAGGCGCAGTACCAAGAGATGAC
87	ProlyspProalysProProalysProleuProvalAlaProglySer-----Ser 107
79	AGCGCTCTTCTAGACTGGGAAAGAAACCTTTCCTGCAGCTCTGATGGACAAAGGCCAC
103	lystHnglvalaIaasPthrglucInProvalLeuthleuProasPleuAlaIuglInPhe
139	ACTGATGATGACTAGATGATGACCCCTGAGCTTGGATGGATGAGAGACATGCGACTTAA
123	AlaProProasPvalAlaProProleu-----Ilelys 134
199	ATTATTACGACCCCGGCTCATTAAGAAATCTGAATGACATACACTATTTCAGGTT
135	LeuLeuGlualIleGlulyselyleuIn---CysserThleuTyArgThGln 155
259	GCAATGGACACTCCCTCCCTTCCCTTACACACCAAGACCTTATCTCATTTGACACCGCACC
154	SetSerSetasPProalaglLeu-----161
319	TGGACACACAGACAGGTTGGAAAGAGTGGCAAAACC---ATTTCAGGAGGTGACA
162	-----ArgInleuLeuaspCysasPProSerValasPleuaspValPhe 177
376	AGCCAAACATTAAAGAGATGCGATCCGTGAAGAAAGAAAGATTCCTTTAACACCTCT
178	AspGlnHisValIleuAlaasPrla---PheIysArgTyIleuAlaasPleuProasPPro
436	CGGCTCTCATTAACCTTCCGAAGAAGTACCAACCCCTTGCCCTCGAGCCGGAAGACAGC
197	-----ValIleProvalAlaValIlyrasGlu 205
496	AGGCCACTTTATCTCAAGACACACACTTTCAGAAAGTCCAGGAAATGCCAGATACATA
206	MetMetsetIleuAlaIn-----GluValIProserSetGluasPtyIle 222
556	ACCTTAAGGAGATTAAAGTGGCTCTTGAAAGCAAAAAAGTTCTCATTAACACAG-----
221	GlnLeu-----LeuTyLysLeuIleArgSerProasInIleProHisGlnTyTyrPleu
610	-----AGAAAGCTGAATACATCATCTG 633
239	ThrlleuGlnTyLeuLeuLysHisPhePheLysLeuSerGlnAlaIasPleuValasPhe 256
634	TTAGAAACCAAAATATCAAGAGATTTCCACATTTAGACAGTCTTCATTCACGACA 693
259	LeuAsnAlaIATGAlaIeLysGlnIle-----PheSerHisValIleuPheArgPhe 273


```

OY 694 AGCACCAGCAGTGGCGAAACACAGATCCATAGA----- 726
Db 276 ProAlaIaSerSerAspAsnThrGluHisLeuIleIysAlaValGluLeuLeuIleSer 295
OY 727 -----GAGGACATGCAGCCCTGTCTCTCAGAGATGCCAGCTCCAGCCAGTGC 777
Db 296 AlaGluTrpSerGluArgGlnProAlaProAlaLeuProProIysProProIysProThr 315
OY 778 AGCCCTCAGCAAAATATACGCGCCCTTAATACACAAGGTGGAGAACACCTTTGCCCAA 837
Db 316 SerIleAlaAsnAsnSerMet----- 322
OY 838 AGGCTGTATGAAAAAGATGTCGACGACAATGAATGGTATCGAATACAGCCGCCAG 897
Db 323 ---AsnAsnMetSerLeuGlnAspArgIaGluTrpTyrGlyAspIleSerArgGlu 341
OY 898 GCATGTGAAAGGCAATTCATGAAAGAGAAACAAAGATGTAGTTTCTGGTCCGAAATTGT 957
Db 342 GluValAsnGluLys--LeuArgAspThrAlaAspGlyThrPheLeuValArgAspAla 360
OY 958 TCCCAAAATTCACAAGGAGAGCCCTATGTTTGGCTGGTATTATCAGAAACAATGTAC 1012
Db 361 SerThrLysMetHisGlyAsp---TyrThrLeuThrLeu-----ArgLysGlyGly 376
OY 1018 AATGTAA---ATCCGCTTCTGGAGAGAGATCAGCACTTGGCCCTGGGAGCAGACTC 1074
Db 377 AsnAsnLysLeuIleIysIlePheHisAspGlyLysTyrGlyPheSerAspProLeu 396
OY 1075 AGAGAGATGAGAAAGTTGATGATAGTAAAGACATGCAATCAGCAATACAGAAAT 1128
Db 397 -----ThrPheAsnSerValValGluLeuIleIleAsnHisTyrArgAsn 410

```

RESULT 7	BTBK_MOUSE	STANDARD;	PRT;	659 AA.
ID	BTBK_MOUSE			
AC	P35991: 061365;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	01-JUN-2002 (Rel. 41, Last annotation update)			
DE	tyrosine-protein kinase BTK (EC 2.7.1.112) (Bruton's tyrosine kinase)			
DE	(Adamaglobulinemia tyrosine kinase) (ATK) (B cell progenitor kinase) (BPK) (kinase EMB).			
GN	BTK OR BPK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93145329; PubMed=8425221;			
RA	Tsukada S., Saitran D.C., Rawlings D.J., Parolinf O., Allen R.C.,			
RA	Kliskar I., Sparkes R.S., Kubagawa H., Mohandas T., Quan S.,			
RA	Belmont J.W., Cooper M.D., Conley M.E., Witte O.N.;			
RT	"Deficient expression of a B cell cytoplasmic tyrosine kinase in			
RL	human x-linked agammaglobulinemia."			
RL	Cell 72:279-290(1993).			
LN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93236578; PubMed=8476425;			
RA	Yanada N., Kawakami Y., Kimura H., Fukumachi H., Baler G.,			
RA	Altman A., Kato T., Inagaki Y., Kawakami T.;			
RT	"Structure and expression of novel protein-tyrosine kinases, Emb and			
RL	Emc, in hematopoietic cells."			
RL	Biochem. Biophys. Res. Commun. 192:231-240(1993).			
LN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95081608; PubMed=7989760;			
RA	Siders P., Mueller S., Shleis H., Jin H., Khan W.N., Nilsson L.,			
RA	Parkinson E., Thomas J.D., Branden L., Larsson I., Paul W.E.,			
RA	Rosen F.S., Alt F.W., Vetric D., Smith C.I.E., Xanthopoulos K.G.;			
RT	"Genomic organization of mouse and human Bruton's agammaglobulinemia			
RL	tyrosine kinase (Btk) loci."			

RL J. Immunol. 153:5607-5617(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C129;
 RX MEDLINE=95352959; PubMed=7626884;
 RA Oeltjen J.C., Liu X., Lu J., Allen R.C., Muzny D.M., Belmont J.W.,
 RA Gibbs R.A.;
 RT "Sixty-nine kilobases of contiguous human genomic sequence containing
 RT the alpha-galactosidase A and Bruton's tyrosine kinase loci.";
 RL Mamm. Genome 6:334-338(1995).
 RN [5]
 RP VARIANT XID CYS-28.
 RX MEDLINE=93324903; PubMed=8332901;
 RA Rawlings D.J., Saffran D.C., Tsukada S., Largaespada D.A.,
 RA Copeland J.C., Cohen L., Mohr R.N., Bazan J.F., Howard M.,
 RA Copeland N.G., Jenkins N.A., Witte O.N.;
 RT "Mutation of unique region of Bruton's tyrosine kinase in
 RT immunodeficient Xid mice.";
 RL Science 261:358-361(1993).
 CC -1. FUNCTION: PLAYS A CRUCIAL ROLE IN B-CELL ONTOGENY.
 CC -1. CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1. SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -1. PFM: Autophosphorylated on Tyr-223 and Tyr-551. The tyrosine
 CC phosphorylation of Tyr-223 may create a docking site for a SH2
 CC containing protein.
 CC -1. DISEASE: DEFECTS IN BTK ARE THE CAUSE OF MURINE X-LINKED
 CC IMMUNODEFICIENCY (XID).
 CC -1. SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
 CC SUBFAMILY.
 CC -1. SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1. SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1. SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, L08967; AAA37316.1; -;
 DR EMBL, L10627; -; NOT_ANNOTATED_CDS.
 DR EMBL, L29788; AAA6943.1; -;
 DR EMBL, U58105; AAB47246.1; -;
 DR HSSP; Q06187; IBWN.
 DR MGD; MGI:88216; Btk.
 DR InterPro: IPR001562; BTK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00065; Pkinase; 1.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00779; BTK; 1.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PRINTS; PR00402; TECBTKDOMAIN.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR ProDom; PD000066; SH3; 1.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00107; BTK; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00219; TYRKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L08967; AAA37316.1; -
DR EMBL; L10627; -; NOT_ANNOTATED_CDS.
DR EMBL; L29788; AAA66943.1; -
DR EMBL; U58105; AAB47246.1; -
DR HSSP; Q06187; IBWN.
DR MGD; MGI:88216; Btk.
DR InterPro; IPR001562; BTK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00402; TECBTDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

```



```

Query Match: 6.93% Indels: 168
DB: 1 Gaps: 18
US-09-856-061-3 (1-1129) x 3BP2_MOUSE (1-559)

QY 166 CTTGGATGGAAGACGACATGGCAGTAAATTTTACACCGCGCCATTAAGAA 225
   |||||
DB 82 MetArgAla1aGluGluThrThrSerAsnValPheProPheLysIleIleHisIle 101

QY 226 TCGAATATGACATACACATTTTCAAGGTTGCATGAC----- 267
   |||||
DB 102 SerLysLysHisArgThrThrPhePheSerAlaSerSerGluAspGluArgLysSer 121

QY 268 -----ACGCCCTCCGCTAGAC 285
   |||||
DB 122 MetAlaPheValArgArgGluIleGluHisPheHisGluLysGluLysGluLeuProLeuAsp 141

QY 286 ACCAGGACCTTATGTCATTTGACAGCCGACCTGGACACACAGACAGAGTTGAAAGA 345
   |||||
DB 142 ThrSerAspSerSer-----AspThrAspSerPheTyrGlyAla 155

QY 346 GTGGACAACCCATTTCCAGGAGCTCAGACGCCAAACATTAAAGAGATGCATCCGTA 405
   |||||
DB 156 ValGluArgProIleAspIleSerLeuSerSerTyrProMetAspAsnGluAspTyrGlu 175

QY 406 AGAAGACAGATTCCT-----TTACCACTCTCGGCT----- 441
   |||||
DB 176 HisGluAspGluAspSerTyrLeuGluProAspSerProGlyProMetLysLeuGlu 195

QY 442 ---CTTATACACTTCGAGAAAGTACCAACCTTCCCTCGAGCCGAGAGACAGCAG 498
   |||||
DB 196 AspAlaLeuThrTyrProProAlaTyrProPro---ProProValProValProArgLys 214

QY 499 CCACCTTATCT-----CAGACACACACCTT----- 525
   |||||
DB 215 ProAlaPheSerAspLeuProAlaHisSerPheThrSerLysSerProSerProLeu 234

QY 526 -----CCAGACTCCAGGGAATGCCAGTCAAGTAAAGCTTA---AGGACTTA 570
   |||||
DB 235 LeuProProProProProLysArgGlyLeuProAspThrGlySerAlaProGluAspAla 254

QY 571 AGTGAAGTCTTGAACACGAAAGTCTCATACACAGAGAGAGCTGATACACTCAT 630
   |||||
DB 255 LysAspAlaLeuGlyLeuArgArgValGluProGlyLeuArgValProAlaThrProArg 274

QY 631 CTGTTGAACAAACCAATACTCAAGAGATTCACCTGCACTTATGAC----- 675
   |||||
DB 275 ArgMetSerAspProProMetSerAsnValProThrValProAsnLeuArgLysHisPro 294

QY 675 ----- 675

DB 295 CysPheArgAspSerValAsnProGlyLeuGluProThrProGlyHisGlyThrSer 314

QY 676 -----AGTTCTCATTCACGACACACACAGTGTGCAAAACGACAGAT----- 720
   |||||
DB 315 SerValSerSerSerThrThrMetAlaValAlaThrSerArgAsnGlySasplLysLeu 334

QY 721 -----CATAGAGAGGACGATGACGCC----- 741
   |||||
DB 335 SerPheHisLeuSerSerArgGlyProProThrSerGluProProProValProAlaAsn 354

QY 742 -----TGTTCTCTCAGAGATGCCAG----- 762
   |||||
DB 355 LysProLysPheLeuLysIleAlaGluProSerProArgGluAlaAlaLysPheAla 374

QY 763 -----CCTCCAGCCAGCTGCACAGCCATCAAGAAATATATCTGCTATAATACACAGC 816
   |||||
DB 375 ProValProProValAlaProArgProProValGluLysMetProMetProGluAlaThr 394

QY 817 TGGAGACACACTTTCCCAAAAGTCTGATAGAAAGATGTCAGAC----- 864
   |||||
DB 395 ValArgProAlaValLeuProArgProGluAsnThrProLeuProHisLeuGluArgSer 414

QY 864 ----- 864
DB 415 ProProAspGlyLysIleSerPheArgGlyPheSerPheGluLysAlaArgGlnProSerGln 434

QY 865 -----AATGAA 870
DB 435 AlaAspThrGlyGluLysAspSerAspGluAspTyrGluLysValProLeuProAsnSer 454

QY 871 TGGTACATTGGAGAAATACAGCCCGCAGCGAGTGAAGAGGACATTGATGAAG----- 921
   |||||
DB 455 ValPheValAsnThrThrGluSerCysGluValGluArgLeuPheLysAlaThrAspPro 474

QY 922 -----GAGAACAGATGATGCTTCTGTCGAGATGTTGTCACAAATCCAGGAA 975
   |||||
DB 475 ArgGlyGluProGlnAspGlyLeuTyrCysIleArgAsnSerSerThrLysSerGlyLys 494

QY 976 GAGCCCTATGTTTGGCTGTGTTATGAG-----AACAACTCTACATGTAAATC 1029
   |||||
DB 495 -----ValLeuValValTrpAspGluSerSerAsnLysValArgGlnTyrArgIle 511

QY 1030 CGCTTCTGAGAGGAAATACAGACTTCCCTCGGACAGAGACTCAGAGAGATGGAAG 1089
   |||||
DB 512 -----PheGluLysAspSerLysPheTyr-----LeuGluGlyGluValLeu 525

QY 1090 TTGATTCAGTAGACATCATCATGACACTAC 1122
   |||||
DB 526 PheAlaSerValGlySerMetValGluHisTyr 536

RESULT 9
BTK_HUMAN
ID BTK_HUMAN STANDARD; PRT; 659 AA.
AC 006187;
DT 01-JUN-1994 (Rel. 29, last sequence update)
DT 01-JUN-1994 (Rel. 41, last annotation update)
DE 15-JUN-2002 (Rel. 41, last annotation update)
DE Tyrosine-protein kinase BTK (EC 2.7.1.112) (Bruton's tyrosine kinase)
DE (Agammaglobulinemia tyrosine kinase) (ATK) (B cell progenitor
DE kinase) (BPK).
GN BTK OR ATK OR AGMX1 OR BPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93140868; PubMed=8380905;
RA Vetric D., Vorechovsky I., Sideras P., Holland J., Davies A.,
RA Flinter F., Hammarstrom L., Kinon C., Levinsky R.J., Bobrow M.,
RA Smith C.I.E., Bentley D.R.;
RT "The gene involved in X-linked agammaglobulinemia is a member of the
RT src family of protein-tyrosine kinases."
RL Nature 361:226-233(1993).
RN [2]
RP ERRATUM.
RA Vetric D., Vorechovsky I., Sideras P., Holland J., Davies A.,
RA Flinter F., Hammarstrom L., Kinon C., Levinsky R.J., Bobrow M.,
RA Smith C.I.E., Bentley D.R.;
RL Nature 364:362-362(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Blood;
RX MEDLINE=94377492; PubMed=8090769;
RA Ohta Y., Haire R.N., Litman R.T., Fu S.M., Nelson R.P., Kratz J.,
RA Kornfeld S.J., la Morena M., Good R.A., Litman G.W.;
RT "Genomic organization and structure of Bruton agammaglobulinemia
RT tyrosine kinase: localization of mutations associated with varied
RT clinical presentations and course in X chromosome-linked
RT agammaglobulinemia."
RL Proc. Natl. Acad. Sci. U.S.A. 91:9062-9066(1994).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=95012452; PubMed=7927535;
RA Rohrer J., Parolini O., Belmont J.W., Conley M.E.;

```

- RT "The genomic structure of human BTK, the defective gene in X-linked
RT agammaglobulinemia.";
RL Immunogenetics 40:319-324(1994).
- [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-95352959; PubMed-7626884;
RA Oeltjen J.C., Iku X., Lu J., Allen R.C., Muzny D.M., Belmont J.W.,
RA Gibbs R.A.;
RT "Sixty-nine kilobases of contiguous human genomic sequence containing
RT the alpha-galactosidase A and Bruton's tyrosine kinase loci.";
RL Mamm. Genome 6:334-338(1995).
- [6]
RP SEQUENCE FROM N.A.
RA Oeltjen J.C., Malley T.M., Muzny D.M., Miller W., Gibbs R.A.,
RA Belmont J.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
- [7]
RP SEQUENCE OF 1-442 FROM N.A.
RX MEDLINE-93145329; PubMed-8425221;
RA Tsukada S., Saffran D.C., Rawlings D.J., Parolini O., Allen R.C.,
RA Kilsak I., Sparks R.S., Kubagawa H., Mohandas T., Quan S.,
RA Belmont J.W., Cooper M.D., Conley M.E., Witte O.N.;
RT "Deficient expression of a B cell cytoplasmic tyrosine kinase in
RT human X-linked agammaglobulinemia.";
RL Cell 72:279-290(1993).
- [8]
RP PHOSPHORYLATION OF GTF2I, AND MUTAGENESIS.
RX MEDLINE-97165069; PubMed-9012831;
RA Yang W., Desiderio S.;
RT "BAP-135, a target for Bruton's tyrosine kinase in response to B cell
RT receptor engagement.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:604-609(1997).
- [9]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-170.
RX MEDLINE-97361821; PubMed-9218782;
RA Hyvoenen M., Saraste M.;
RT "Structure of the PH domain and Btk motif from Bruton's tyrosine
RT kinase: molecular explanations for X-linked agammaglobulinemia.";
RL EMBO J. 16:3396-3404(1997).
- [10]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-170.
RX MEDLINE-99216540; PubMed-10196129;
RA Baraldi E., Carugo K.D., Hyvoenen M., Surdo P.L., Riley A.M.,
RA Potter B.V.L., O'Brien R., Ladbury J.E., Saraste M.;
RT "Structure of the PH domain from Bruton's tyrosine kinase in complex
RT with inositol 1,3,4,5-tetrakisphosphate.";
RL Structure 7:449-460(1999).
- [11]
RP STRUCTURE BY NMR OF 209-275.
RX MEDLINE-98153176; PubMed-9485443;
RA Hansson H., Mattsson P.T., Allard P., Haapaniemi P., Vihinen M.,
RA Smith C.I.E., Heerd T.;
RT "Solution structure of the SH3 domain from Bruton's tyrosine kinase.";
RL Biochemistry 37:2912-2924(1998).
- [12]
RP REVIEW ON XLA VARIANTS.
RX MEDLINE-96174621; PubMed-8594569;
RA Vihinen M., Iwata T., Kinon C., Kwan S.-P., Ochs H.D.,
RA Vorechovsky I., Smith C.I.E.;
RT "BTKbase, mutation database for X-linked agammaglobulinemia (XLA).";
RL Nucleic Acids Res. 24:160-165(1996).
- [13]
RP REVIEW ON XLA VARIANTS.
RX MEDLINE-97169387; PubMed-9016530;
RA Vihinen M., Belohradsky B.H., Halre R.N., Hollinski-Peder E.,
RA Kwan S.-P., Lappalainen I., Lehtvaesalho H., Lester T., Melndl A.,
RA Ochs H.D., Ollila J., Vorechovsky I., Weiss S., Smith C.I.E.;
RT "BTKbase, mutation database for X-linked agammaglobulinemia (XLA).";
RL Nucleic Acids Res. 25:166-171(1997).
- [14]
RP VARIANTS XLA TRP-288; GLY-307; ASP-607 AND S-V-F-S-T-R-103 INS.
RX MEDLINE-94214481; PubMed-8162056;
RA Bradley L.A.D., Sweatman A.K., Lovering R.C., Jones A.M., Morgan G.,
RA
- RA Levinsky R.J., Kinon C.;
RT "Mutation detection in the X-linked agammaglobulinemia gene, BTK,
RT using single strand conformation polymorphism analysis.";
RL Hum. Mol. Genet. 3:79-83(1994).
- [15]
RP VARIANTS XLA HIS-28 AND TRP-288.
RX MEDLINE-94214435; PubMed-8162018;
RA de Weers M., Mensink R.G.J., Kraakman M.E.M., Schuurman R.K.B.,
RA Hendriks R.W.;
RT "Mutation analysis of the Bruton's tyrosine kinase gene in X-linked
RT agammaglobulinemia: identification of a mutation which affects the
RT same codon as is altered in immunodeficient xid mice.";
RL Hum. Mol. Genet. 3:161-166(1994).
- [16]
RP SEQUENCE FROM N.A., AND VARIANTS XLA S-334, R-506, Q-520, W-562 AND
RP K-630.
RX MEDLINE-95152493; PubMed-7880320;
RA Hagemann T.L., Chen Y., Rosen F.S., Kwan S.-P.;
RT "Genomic organization of the Btk gene and exon scanning for mutations
RT in patients with X-linked agammaglobulinemia.";
RL Hum. Mol. Genet. 3:1743-1749(1994).
- [17]
RP VARIANTS XLA D-113; C-361; Q-520; P-542; W-562; K-630 AND P-652.
RX MEDLINE-95152494; PubMed-7849697;
RA Conley M.E., Fitch-Hilgenberg M.E., Cleveland J.L., Parolini O.,
RA Rohrer J.;
RT "Screening of genomic DNA to identify mutations in the gene for
RT Bruton's tyrosine kinase.";
RL Hum. Mol. Genet. 3:1751-1756(1994).
- [18]
RP VARIANTS XLA H-28; P-33; P-408; G-589; D-613 AND 260-Q--E-280 DEL.
RX MEDLINE-95152522; PubMed-7849721;
RA Zhu Q., Zhang M., Winkelsstein J., Chen S.-H., Ochs H.D.;
RT "Unique mutations of Bruton's tyrosine kinase in fourteen unrelated
RT X-linked agammaglobulinemia families.";
RL Hum. Mol. Genet. 3:1899-1900(1994).
- [19]
RP VARIANTS XLA E-430; Q-520; Q-525; P-562; V-582; G-589; E-594 AND
RP D-613.
RX MEDLINE-95108046; PubMed-7809124;
RA Vihinen M., Vetric H.S., Maniar H.S., Ochs H.D., Zhu Q., Vorechovsky I.,
RA Webster A.D.B., Notarangelo L.D., Nilsson L., Sovadski J.M.,
RA Smith C.I.E.;
RT "Structural basis for chromosome X-linked agammaglobulinemia: a
RT tyrosine kinase disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12803-12807(1994).
- [20]
RP VARIANT XLA PHE-64, AND CHARACTERIZATION OF OTHER XLA VARIANTS.
RX MEDLINE-95151728; PubMed-7849006;
RA Vihinen M., Zvelebil J.J.M., Zhu Q., Broolmans R.A., Ochs H.D.,
RA Zegers B.J.M., Nilsson L., Waterfield M.D., Smith C.I.E.;
RT "Structural basis for pleckstrin homology domain mutations in
RT X-linked agammaglobulinemia.";
RL Biochemistry 34:1475-1481(1995).
- [21]
RP VARIANTS XLA S-25; W-288; M-370; V-509; P-525; K-526; V-582
RP AND R-594.
RX MEDLINE-95227177; PubMed-7711734;
RA Vorechovsky I., Vihinen M., de Saint Basile G., Honsova S.,
RA Hammarstrom L., Mueller S., Nilsson L., Fischer A., Smith C.I.E.;
RT "DNA-based mutation analysis of Bruton's tyrosine kinase gene in
RT patients with X-linked agammaglobulinemia.";
RL Hum. Mol. Genet. 4:51-58(1995).
- [22]
RP VARIANTS XLA IYS-567; LEU-587 AND HIS-641.
RX MEDLINE-95359977; PubMed-7633420;
RA Jin H., Webster A.D.B., Vihinen M., Sideras P., Vorechovsky I.,
RA Hammarstrom L., Bernatowska-Matuszkiewicz E., Smith C.I.E.,
RA Bobrow M., Vetric D.;
RT "Identification of Btk mutations in 20 unrelated patients with
RT X-linked agammaglobulinemia (XLA).";
RL Hum. Mol. Genet. 4:693-700(1995).
- [23]

Best Local Similarity: 20.55% Mismatches: 172
 Query Match: 6.73% Indels: 142
 DB: 1 Gaps: 20

US-09-856-061-3 (1-1129) x 3BP2_HUMAN (1-561)

OY 22 AAAAAAGGTCATGCGCTGCATGATAGTCCACAGGCCAGTACCAGAGATGAACAAG 81
 DB 104 LysHisAlaArgThrTrp-----PhePheSerAlaSerSerGluGluArg----- 118
 OY 82 CCTCTTCAGACTGGGAAAGAACTTGTGTCAGCTCCGATGAGAGCAAAAGCCACAGT 141
 DB 119 -----LysSerTrpMetAlaLeuLeuAlaArgGluIleGluSerPhe 132
 OY 142 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 201
 DB 133 HisGluLysLysAlaPheLeuPheLeuAspThrSerAspSerSerSerAspThrSerPhe 152
 OY 202 TTACACAGCC-----CGGCCTATA-----AAG 222
 DB 153 TyrGluAlaValAlaGluArgProValAspIleSerLeuSerProTyrProThrAspAsnGlu 172
 OY 223 GAATGTGATATGATGATATGATGATATGATGATGATGATGATGATGATGATGAT 282
 DB 173 AspTyrGluLysAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 191
 OY 283 GACACCCAGGACCTCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 342
 DB 192 ---GluArgLeuGluAspAlaLeuMetHisProProAlaTyrProProProProValPro 210
 OY 343 AGAGTGGACAAACCCATTTCCAGGAGCTC-----AGAAGCCAAAACATTAAAGA 393
 DB 211 ThrProArgLysProAlaPheSerAspMetProArgAlaHisSerPheThrSerLysGlu 230
 OY 394 GATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
 DB 231 ProGlyProLeu-----LeuProProProProProProProProProProPro 246
 OY 439 -----CCTCTCATTAACCTTCCGGAAGATAC 465
 DB 247 ValGluLeuAlaAlaGluAspSerLysAspArgProLeu-----CysProArgArgAla 264
 OY 466 CAACCTTGCCCTGAGCCGCGAGAGCAGCAGC-----CCACCTTATATCAG 513
 DB 265 GluProCysProArgValAlaThrProArgArgMetSerAspProProLeuSer--- 283
 OY 514 AGACACACCTTCCAGAGTCCAGGGAATG----- 543
 DB 284 -----ThrMetProThrAlaProGlyLeuArgLysProProCysPheArgGluSerAla 301
 OY 544 -----CCAGTCCAGATTAAGCTTAAGGAGCTTAAGTGAAGTGC 579
 DB 302 SerProSerProGluProTrpTrpTrpProGluHisGluAlaCysSerThrSerSerAla 321
 OY 580 CTTGAGCAGGAAAAAGTTCTCATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 639
 DB 322 IleMetAlaThrAlaThrSerArgAsnGlyLysLysSerPheHisLeuSerPro 341
 OY 640 AAC-----CAAAATACTCAAGAGATTCACCTTGCATTAGACAGTTCTTATCAGCAG 693
 DB 342 ArgGlyProProTrpThrSerGluProProProValProAlaAsnLysProLysPhe----- 359
 OY 694 AGCAACACAGTGTGCAAAAGAGATCATAGAGGAGCAGCAGCCCTGTTCTCTCAG 753
 DB 360 -----LeuLysIleAlaGluLysAspProProArgGluAlaAlaMetProGly 375
 OY 754 AGATGCCAGCCCTCAGCCAGCCAGCCCTCAGAAATATATGCTTAATAATACAG 813
 DB 376 LeuPheValProProValAlaProArgProProAlaLeuLysLeuProValProGluAla 395
 OY 814 AGCTGAGACACGCTTCCCAAAAGGCTGTATAGAAAAGATGTCAGAC----- 864
 DB 396 MetaLysProAlaValLeuProArgProGluLysProGluLysProGluHisLeuGlnArg 415

OY 864 ----- 864
 DB 416 SerProProAspGlyGlnSerPheArgSerPheSerPheGluLysProArgGlnProSer 435
 OY 865 -----AAT 867
 DB 436 GlnAlaAspThrGlyGlyAspAspSerAspGluAspTyrGluLysValProLeuProAsn 455
 OY 868 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 921
 DB 456 SerValPheValAsnThrThrGluSerCysGluValGluArgLeuPheLysAlaThrSer 475
 OY 922 -----GAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 972
 DB 476 ProArgGlyGluProGlnAspLysLeuTyrCysIleArgAsnSerSerThrLysSerGly 495
 OY 973 GAGAGCCCTATGTTTGGCTGTGTTTATGAG-----ACAAGTCTACATGTAA 1026
 DB 496 Lys-----ValLeuValValTrpAspGluThrSerAsnLysValArgAsnTyrArg 512
 OY 1027 ATCCGCTTCCGTGAGAGGATGACAGCTTGGCCCGGAGACGACTCAGAGAGATGAG 1086
 DB 513 Ile-----PheGluLysAspSerLysPheTyr-----LeuGluGlyGluVal 526
 OY 1087 AAGTTGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1122
 DB 527 LeuPheValSerValGlySerMetValGluHisTyr 538

RESULT 12
 P85B_BOVIN STANDARD: PRT; 724 AA..
 ID P23726;
 AC 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) [PtdIns-3-kinase p85-beta].
 GN PIK3R2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91191567; PubMed=1707345;
 RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Larrea F., Panayotou G.,
 RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
 RA Courtenidge S.A., Parker P.J., Waterfield M.D.;
 RA "Characterization of two 85 kd proteins that associate with receptor
 RA tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase.";
 RA Cell 65:91-104(1991).
 RL [2]
 RN CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.
 RP MEDLINE=93049176; PubMed=1330535;
 RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,
 RA Fry M.J., Blundell T.L., Wolmer A., Waterfield M.D.;
 RA "Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
 RA domain with a PDGF receptor phosphorylation site: structural features
 RA and analysis of conformational changes.";
 RL EMBD J. 11:4261-4272(1992).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES THROUGH ITS SH2 DOMAIN AND REGULATES THEIR KINASE
 CC ACTIVITY. DURING INSULIN STIMULATION, IT ALSO BINDS TO IRS-1.
 CC -1- SUBUNIT: HETEROIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN BRAIN AND TESTIS. LOWER
 CC LEVELS IN ADIPOSE TISSUE, KIDNEY, HEART, LUNG AND SKLETERAL MUSCLE.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: HIGH, WITH OTHER P85 SUBUNITS, AND WITH P85-ALPHA AND
 CC P85-BETA SUBUNITS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D88532; BAA13636.1; -
 CC EMBL: AF028785; AAC39696.1; -
 CC HSSP: P23727; 2PMB.
 CC Genew: HGNC:8981; PIK3R3.
 CC MIM: 606076; -
 CC InterPro: IPR001720; PI3kinase_p85.
 CC InterPro: IPR000980; SH2.
 CC Pfam: PF00017; SH2; 2.
 CC PRINTS: PRO0678; PI3KINSEP85.
 CC PRINTS: PR00401; SH2DOMAIN.
 CC PRODOM: PD000093; SH2; 2.
 CC SMART: SM00252; SH2; 2.
 CC PROSITE: PS50001; SH2; 2.
 CC SH2 domain; Repeat; Phosphorylation; Alternative splicing.
 CC KW DOMAIN 34 44
 CC FT DOMAIN 34 44
 CC FT DOMAIN 65 160
 CC FT MOD_RES 358 452
 CC FT MOD_RES 361 341
 CC FT VARSPITC 256 314
 CC FT VARSPITC 21 21
 CC FT CONFLICT 21 21
 CC FT P -> L (IN REF. 2).
 CC SO SEQUENCE 461 AA; 54462 MW; C03ECFF22BEB5089 CRC64;
 CC -----
 CC Alignment Scores:
 CC Pred. No.: 0.0308 Length: 461
 CC Score: 128.00 Matches: 49
 CC Percent Similarity: 45.64% Conservative: 19
 CC Best Local Similarity: 32.89% Mismatches: 45
 CC Query Match: 6.34% Indels: 36
 CC DB: 1 Gaps: 10
 CC -----
 CC US-09-856-061-3 (1-1129) x P55G_HUMAN (1-461)
 CC QY 718 GATCATAGAGAGGAGCGATCGACCCCTGTCTCTCTCTAG-----AGATGCCAG 762
 CC DB 14 AsPTpArgGLuValMetMetProTyrSerThrGluLeuIlePheTyrIleGluMetasp 33
 CC QY 763 CCTCAGCAGCAGCTGCAGCCCTCAGCAAAATATACCTGCCCCCTATTAATACCAAGCTGAGAGA 822
 CC DB 34 PropProlA-----LeuPropProlYs----- 40
 CC QY 823 CCACCTTCCCAAAAGAGCT-----GATAGAAGAGAT-----GTC 858
 CC DB 41 PropProlYsPrometThSerAlaValProAsnGlyMetLysAspSerSerValSerLeu 60
 CC QY 859 CAGCAAAATGATGTATGATGAGATATACAGCCCGCAGGAGCGAGGAGGATTCATCATG 918
 CC DB 61 GluAspAlaGluTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyr 79
 CC QY 919 AAGGAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 978
 CC DB 80 ArgAspMetProAspGlyThrPheLeuValArgAspAlaSerThrLysMetGlnGlyasp 99

QY 979 CCTATGCTTTGGCGTGTGTTTATGAGAACAAAGTCTNCAATGTAA--ATCCGCTTC 1035
 DB 100 ---TyrThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIle 115
 QY 1036 CTGAGAGAGCAATCAGCAGCTTCCCTGGGAGACAGCTCAGAGAGATGAGAAGTTTGAT 1095
 DB 116 TyrHisArgAspGlyLysTyrGlyLysPheSerAProLeu-----ThrPheAsn 131
 QY 1096 TCAGTGAAGACATCATCATCAACACTAC 1122
 DB 132 SerValValGluLeuIleAsnHisTyr 140
 CC -----
 CC RESULT 14
 CC P85_MOUSE STANDARD; PRT; 722 AA.
 CC ID P85_MOUSE
 CC AC 008908;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 CC P85-beta subunit) (Ptdins-3-kinase p85-beta).
 CC GN PIK3R2.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_Taxid=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=NIH SWISS;
 CC RX MEDLINE=98241181; PubMed=9582025;
 CC RA Janssen J.W.G., Schleithoff L., Bartam C.R., Schulz A.S.;
 CC RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase
 CC RT p85beta subunit and HUNKOR8, a putative deubiquitinating enzyme.";
 CC RL Oncogene 16:1767-1772(1998).
 CC CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC -1- SUBUNIT: HETEROIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y13569; CAA73903.1; -
 CC HSSP: P23727; 2PMB.
 CC MGD: MGI:1098772; PIK3R2.
 CC InterPro: IPR001720; PI3kinase_p85.
 CC InterPro: IPR000198; RHO-GAP.
 CC InterPro: IPR000980; SH2.
 CC InterPro: IPR001452; SH3.
 CC Pfam: PF00017; SH2; 2.
 CC Pfam: PF00018; SH3; 1.
 CC Pfam: PF00620; RHO-GAP; 1.
 CC PRINTS: PRO0678; PI3KINSEP85.
 CC PRINTS: PR00401; SH2DOMAIN.
 CC PRODOM: PD000066; SH3; 1.
 CC PRODOM: PD000093; SH2; 2.
 CC SMART: SM00324; RHO-GAP; 1.
 CC SMART: SM00252; SH2; 2.
 CC SMART: SM00326; SH3; 1.
 CC PROSITE: PS50001; SH2; 2.
 CC PROSITE: PS50002; SH3; 1.
 CC SH3 domain; SH2 domain; Repeat.

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirsky R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster";
CC Science 287:2185-2195(2000).
CC -1- FUNCTION: ADAPTER PROTEIN WHICH INTERACTS WITH C-TERMINAL PORTION
CC OF MRC, HOMOLOG OF HUMAN DOCK180. MAY PLAY A ROLE IN CELLULAR
CC PROCESSES THROUGHOUT DEVELOPMENT.
CC -1- TISSUE SPECIFICITY: EMBRYONIC ZYGOTIC EXPRESSION IS SEEN IN
CC INVAGINATING PRESUMPTIVE MESODERM AND ECTODERMALLY DERIVED TISSUES
CC DURING GASTRULATION. AT STAGE 8, EXPRESSION IS ALSO SEEN IN
CC ANTERIOR AND POSTERIOR MIDGUT AND CEPHALIC FURROW. BY STAGE 9,
CC EXPRESSION IS HIGHEST IN VISCERAL MESODERM OF ANTERIOR AND
CC POSTERIOR MIDGUT, VENTRAL NERVE CORD AND SOMATIC MESODERM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY
CC THROUGHOUT EMBRYOGENESIS, DECLINES DURING LARVAL STAGES AND
CC REAPPEARS DURING PUPATION.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003844; AAF59362.1; -
DR EMBL: AF112976; AAD28428.1; -
DR HSSP: O64010; ICKA.
DR FlyBase: FBgn002481; Crk.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PRO0018; SH3; 1.
DR PRINTS: PRO0401; SH2DOMAIN.
DR PRINTS: PRO0452; SH3DOMAIN.
DR PRODOM: PD000066; SH3; 1.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00326; SH3; 2.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
KW SH2 domain; SH3 domain; Repeat; Developmental protein.
FT DOMAIN 12 114 SH2.
FT DOMAIN 117 165 SH3 1.
FT DOMAIN 220 259 SH3 2.
SQ SEQUENCE 271 AA; 31205 MW; D1B4FE43150932DC CRC64;

Alignment Scores:
Pred. No.: 0.0342 Length: 271
Score: 127.00 Matches: 33
Percent Similarity: 50.00% Conservative: 14
Best Local Similarity: 35.11% Mismatches: 37
Query Match: 6.29% Indels: 10
DB: 1 Gaps: 3

US-09-856-061-3 (1-1129) x CRK_DROME (1-271)

QY 841 TCTGATAGAAAGATGTCACACATGATGTCATGATGATATACAGCCGCGACGCA 900
DB 7 Serasparg-----AsnseritryipheglyPrometserarglnasp 21

QY 901 GTGAGAGAGCCATTCATGAGAGAGAGAGAGAGATGATGTTCTTGTCGAGATGTTCC 960
DB 22 AlathrgluValleuMetAsnGluArgGluArgGlyValPheLeuValArgAspSerAsn 41
QY 961 ACAAAATCCAG 1020
DB 42 SerllealaGlyasp-----TyrValleuGlyValArgGluAspPrnLysValSerAsn 59
QY 1021 GTAAAAATCCGCTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 60 TyrilleaAsnLysValGlnGlnGlnAspGlnLysValTyrArglle-----Gly 76
QY 1081 GATGAG 1122
DB 77 AspGlnserPheAspAsnLeuProlLysLeuLeuThrPheTyr 90

Search completed: April 21, 2003, 12:22:40
Job time : 30.2418 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:22:47 ; Search time 30.5993 Seconds
(without alignments)
1637.367 Million cell updates/sec

Title: US-09-856-061-4

Perfect score: 2020

Sequence: 1 FGNFSLPKNRSWPRINSATG.....GLRGDEKFDSEVDIEHYKN 376

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2020	100.0	376	22	AAAB81071
2	2009	99.5	428	23	AAE22609
3	2009	99.5	443	23	AAE22608
4	1694	83.9	353	23	AAE22610
5	1132	56.0	435	22	AAAB81070
6	349	17.3	84	21	AAAG03152
7	289	14.3	449	19	AAW59866
8	289	14.3	456	22	AAAB93406
9	286	14.2	456	19	AAW59865
10	273	13.5	457	19	AAW59867

11	238	11.8	46	22	ABB38147	Peptide #5653 enco
12	238	11.8	46	22	ABB23354	Protein #5353 enco
13	238	11.8	46	22	AAW58783	Human brain expres
14	238	11.8	46	22	AAW71296	Human bone marrow
15	238	11.8	46	22	AAW18973	Peptide #5407 enco
16	238	11.8	46	22	AAW31575	Peptide #5612 enco
17	238	11.8	46	23	ABG41096	Human peptide enco
18	236	11.7	95	20	AAW70587	Human SLP-76 SH2 d
19	214	10.6	40	22	ABB39614	Peptide #7120 enco
20	214	10.6	40	22	AAW60320	Human bone marrow
21	214	10.6	40	22	AAW72953	Human brain expres
22	214	10.6	40	22	AAW33179	Peptide #7216 enco
23	214	10.6	40	23	ABG42791	Human peptide enco
24	170	8.3	503	22	ABG22854	Novel human diago
25	167	8.3	945	22	ABB66077	Drosophila melanog
26	147.5	7.3	759	22	ABG21983	Novel human diago
27	147	7.3	724	16	AAW85784	Human GRB-1. Homo
28	147	7.3	724	23	ABG60297	Mouse lymphoma ass
29	146.5	7.3	724	13	AAW26060	Growth Factor Rece
30	146.5	7.3	724	21	AAV97143	Human PI3K p85 pol
31	146.5	7.3	724	22	AAW67442	Amino acid sequenc
32	146.5	7.3	724	22	AAW67821	Human lymphoma ass
33	146.5	7.3	724	23	ABG60298	p85alpha-green flo
34	146.5	7.3	968	19	AAW85011	p85alpha-green flo
35	146.5	7.3	970	19	AAW85020	Human C11A intera
36	144.5	7.2	945	20	AAV06119	Mouse haematopoiet
37	143	7.1	659	17	AAW06708	Mouse Bruton's tyr
38	143	7.1	659	23	ABB76488	Mouse Bruton's tyr
39	142	7.0	611	22	ABG04956	Novel human diago
40	140	6.9	635	22	AAU09009	Human Bruton's tyr
41	140	6.9	659	17	AAW94534	BRK tyrosine kinas
42	140	6.9	659	23	ABB76487	Human Bruton's tyr
43	137	6.8	442	17	AAW06709	Human haematopoiet
44	136.5	6.8	841	22	ABB67384	Drosophila melanog
45	136.5	6.8	841	22	AAW59217	D.melanogaster cor

ALIGNMENTS

RESULT 1

AAAB81071 standard; Protein; 376 AA.

AAAB81071:

25-JUN-2001 (first entry)

Human mast cell-specific immunoreceptor signal transducer.

Human mast cell-specific immunoreceptor signal transducer.

Mast cell; signal transduction; human; allergic disease; MIST;

Mast cell-specific immunoreceptor signal transducer.

Homo sapiens.

JP3146204-B1.

12-MAR-2001.

17-SEP-1999; 99JP-0263778.

17-SEP-1999; 99JP-0263778.

17-SEP-1999; 99JP-0263778.

(KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.

WPI; 2001-310022/33.

N-PSDB; AAF86140.

Mast cell-specific signal-transduction molecule; useful for screening

therapeutic compounds for treating allergies, is specifically expressed

by mouse mast cell -

Example 1; Page 10-11; 12pp; Japanese.

XX This sequence represents human MIST (mast cell-specific immunoreceptor
 CC signal transducer). The invention relates to cDNA encoding a murine mast
 CC cell-specific signal transduction protein. Included in the invention are
 CC cDNA and protein sequences of the mast cell-specific signal transduction
 CC molecule and an expression vector containing the polynucleotide sequence.
 CC The coding sequence of the signal transduction protein can be used for
 CC screening therapeutic compounds which will be useful for treating
 CC allergic diseases.

XX Sequence 376 AA:

Query Match 100.0%; Score 2020; DB 22; Length 376;
 Best Local Similarity 100.0%; Pred. No. 2, 2e-172;

Matches 376; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQNSLPRNRSMPRINSATGQYQRMKPLDMERNFAVLGAGSHSDDDYDDELAMEE 60
 DB 1 FQNSLPRNRSMPRINSATGQYQRMKPLDMERNFAVLGAGSHSDDDYDDELAMEE 60
 QY 61 TWOSIKILPAPRIKESVADTHYFKVAMDTPLDTRTSISIGOPTWNTQRLERVDKPI 120
 DB 61 TWOSIKILPAPRIKESVADTHYFKVAMDTPLDTRTSISIGOPTWNTQRLERVDKPI 120
 QY 121 SRDVRSONIKGDASYRKNKILPPEPRPLITLPRKYQPLPPEPSSRPLSQRHFFPEVOG 180
 DB 121 SRDVRSONIKGDASYRKNKILPPEPRPLITLPRKYQPLPPEPSSRPLSQRHFFPEVOG 180
 QY 181 MFSQSLDLSEVLEAEVPHNQKREESTHLENONTQETPLAISSSFTSNHSVQNRD 240
 DB 181 MFSQSLDLSEVLEAEVPHNQKREESTHLENONTQETPLAISSSFTSNHSVQNRD 240
 QY 241 HRGGMQPCSPQRCOPASCSPHENTLPYKYSWRRPFKRSRDKVDQVQNHENYIGEYSRQA 300
 DB 241 HRGGMQPCSPQRCOPASCSPHENTLPYKYSWRRPFKRSRDKVDQVQNHENYIGEYSRQA 300
 QY 301 VEEAFMKENKDGSLVRCSTKSKEEYVLAFFYENKYNKIRPLENQOFPALGTGLRG 360
 DB 301 VEEAFMKENKDGSLVRCSTKSKEEYVLAFFYENKYNKIRPLENQOFPALGTGLRG 360
 QY 361 DEKPDVEDIIEHYKN 376
 DB 361 DEKPDVEDIIEHYKN 376

RESULT 2
 AAE22609

ID AAE22609 standard; Protein: 428 AA.

AC AAE22609;

DT 26-JUL-2002 (first entry)

DE Human MIST splice variant protein from clone #7.

XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
 KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.

OS Homo sapiens.

PN WO200226986-A2.

PD 04-APR-2002.

PF 28-SEP-2001; 2001WO-US30593.

PR 29-SEP-2000; 2000US-237030P.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB,
 FI
 XX

DR WPI: 2002-372126/40.
 DR N-PSDB: AAD35801.

XX New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
 PS Claim 11; Fig 5; 174pp; English.

XX The present invention relates to novel mast cell immunoreceptor signal
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.
 CC MIST sequences of the invention are useful for preventing, treating or
 CC ameliorating a medical condition in mammalian subject. They are useful
 CC for treating an immune disorder involving hyperactivity of B- or T-
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
 CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
 CC as targets for therapeutic intervention in immune cell disorders and
 CC inflammatory indications, for diagnosis and/or screening of disorders
 CC or diseases associated with expression of MIST, for screening for
 CC antagonists or inhibitors of the interaction of MIST with cellular
 CC signalling components. They are used in assays that detect activation
 CC or induction of various B and T-cell-related neoplasms or cancers.
 CC Sequences of the invention are also used in gene therapy. The present
 CC sequence is human MIST splice variant protein from clone #7.

SQ Sequence 428 AA:

Query Match 99.5%; Score 2009; DB 23; Length 428;
 Best Local Similarity 99.5%; Pred. No. 2, 6e-171;

Matches 374; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQNSLPRNRSMPRINSATGQYQRMKPLDMERNFAVLGAGSHSDDDYDDELAMEE 60
 DB 19 FQNSLPRNRSMPRINSATGQYQRMKPLDMERNFAVLGAGSHSDDDYDDELAMEE 78
 QY 61 TWOSIKILPAPRIKESVADTHYFKVAMDTPLDTRTSISIGOPTWNTQRLERVDKPI 120
 DB 79 TWOSIKILPAPRIKESVADTHYFKVAMDTPLDTRTSISIGOPTWNTQRLERVDKPI 138
 QY 121 SRDVRSONIKGDASYRKNKILPPEPRPLITLPRKYQPLPPEPSSRPLSQRHFFPEVOG 180
 DB 139 SRDVRSONIKGDASYRKNKILPPEPRPLITLPRKYQPLPPEPSSRPLSQRHFFPEVOG 198
 QY 181 MFSQSLDLSEVLEAEVPHNQKREESTHLENONTQETPLAISSSFTSNHSVQNRD 240
 DB 199 MFSQSLDLSEVLEAEVPHNQKREESTHLENONTQETPLAISSSFTSNHSVQNRD 258
 QY 241 HRGGMQPCSPQRCOPASCSPHENTLPYKYSWRRPFKRSRDKVDQVQNHENYIGEYSRQA 300
 DB 259 HRGGMQPCSPQRCOPASCSPHENTLPYKYSWRRPFKRSRDKVDQVQNHENYIGEYSRQA 318
 QY 301 VEEAFMKENKDGSLVRCSTKSKEEYVLAFFYENKYNKIRPLENQOFPALGTGLRG 360
 DB 319 VEEAFMKENKDGSLVRCSTKSKEEYVLAFFYENKYNKIRPLENQOFPALGTGLRG 378
 QY 361 DEKPDVEDIIEHYKN 376
 DB 379 DEKPDVEDIIEHYKN 394

RESULT 3
 AAE22608

ID AAE22608 standard; Protein: 443 AA.

AC AAE22608;

DT 26-JUL-2002 (first entry)

DE Human MIST protein #1.

XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
 KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.

```
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 84
XX FT Modified-site 111
XX FT Modified-site 111
XX FT Domain /label= Tyrosine-phosphorylation_binding_site
XX FT Domain 306..311
XX FT Domain /label= SH3-binding_proline-rich_motif
XX FT Domain 324..407
XX FT Domain /label= SH2_domain
XX PN WO200226986-A2.
XX PD 04-APR-2002.
XX PF 28-SEP-2001; 2001WO-US30593.
XX PR 29-SEP-2000; 2000US-237030P.
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX DR WPI: 2002-372126/40.
XX DR N-PSDB: AAD35800.
XX PT New isolated mast cell immunoreceptor signal transducer polypeptide,
XX PT useful for treating immune disorder involving hyperactivity of B- or
XX PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX PS Claim 11; Fig 3; 171pp; English.
XX CC The present invention relates to novel mast cell immunoreceptor signal
XX CC transducer (MIST) proteins and polynucleotides encoding such proteins.
XX CC MIST sequences of the invention are useful for preventing, treating or
XX CC ameliorating a medical condition in mammalian subject. They are useful
XX CC for treating an immune disorder involving hyperactivity of B- or T-
XX CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX CC lymphoma, tumor or thymoma in a mammal. MIST sequences are useful
XX CC as targets for therapeutic intervention in immune cell disorders and
XX CC inflammatory indications, for diagnosis and/or screening of disorders
XX CC or diseases associated with expression of MIST, for screening for
XX CC antagonists or inhibitors of the interaction of MIST with cellular
XX CC signaling components. They are used in assays that detect activation
XX CC or induction of various B and T-cell-related neoplasms or cancers.
XX CC Sequences of the invention are also used in gene therapy. The present
XX CC sequence is human MIST protein.
XX SQ Sequence 443 AA;
XX Query Match 99.5%; Score 2009; DB 23; Length 443;
XX Best Local Similarity 99.5%; Pred. No. 2.7e-171;
XX Matches 374; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 FQNFSLPKRNSMRINSATGQYORANKPLLDWERNFAVLDGAKGSHDDYDDELMEE 60
DB 34 FQNFSLPKRNSMRINSATGQYORANKPLLDWERNFAVLDGAKGSHDDYDDELMEE 93
OY 61 TWSITILPAPRIKESYADTHYFKVAMDPLDPTSTISIGOPTWNTOTRLERVDKPI 120
DB 94 TWSITILPAPRIKESYADTHYFKVAMDPLDPTSTISIGOPTWNTOTRLERVDKPI 153
OY 121 SRDVSQNIKGDAVSRKKNILPDPRLTLPPKYOPLPEPSSRPPLSQRHTFPEVQG 180
DB 154 SKDVSQNIKGDAVSRKKNILPDPRLTLPPKYOPLPEPSSRPPLSQRHTFPEVQR 213
OY 181 MFSQISIRDLSEVLEAEKVHNRKPESTHLLNQNTQETPLAISSTFTSHNSVONRD 240
DB 214 MFSQISIRDLSEVLEAEKVHNRKPESTHLLNQNTQETPLAISSTFTSHNSVONRD 273
OY 241 HRGGMQPCSPQRCOPASCSPHENILPYKYSWRPPPKSRDKRDVQHNEMWIGEYSROA 300
```

```
DB 274 HRGGMQPCSPQRCOPASCSPHENILPYKYSWRPPPKSRDKRDVQHNEMWIGEYSROA 333
OY 301 VEEAFMKENKDGSEFLVRCSTKSEEPYVLAVYEKNVYVWKTRFLERNOQFALGTGLRG 360
DB 334 VEEAFMKENKDGSEFLVRCSTKSEEPYVLAVYEKNVYVWKTRFLERNOQFALGTGLRG 393
OY 361 DEKFDSEVEDIEEHYKN 376
DB 394 DEKFDSEVEDIEEHYKN 409
RESULT 4
AAE22610
ID AAE22610 standard; Protein: 353 AA.
XX AC AAE22610;
XX AC AAE22610;
XX DT 26-JUL-2002 (first entry)
XX DE Human MIST splice variant protein from clone #12.
XX KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
XX KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
XX KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX OS Homo sapiens.
XX PN WO200226986-A2.
XX PD 04-APR-2002.
XX PF 28-SEP-2001; 2001WO-US30593.
XX PR 29-SEP-2000; 2000US-237030P.
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX DR WPI: 2002-372126/40.
XX DR N-PSDB: AAD35802.
XX PT New isolated mast cell immunoreceptor signal transducer polypeptide,
XX PT useful for treating immune disorder involving hyperactivity of B- or
XX PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX PS Claim 11; Fig 8; 171pp; English.
XX CC The present invention relates to novel mast cell immunoreceptor signal
XX CC transducer (MIST) proteins and polynucleotides encoding such proteins.
XX CC MIST sequences of the invention are useful for preventing, treating or
XX CC ameliorating a medical condition in mammalian subject. They are useful
XX CC for treating an immune disorder involving hyperactivity of B- or T-
XX CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
XX CC as targets for therapeutic intervention in immune cell disorders and
XX CC inflammatory indications, for diagnosis and/or screening of disorders
XX CC or diseases associated with expression of MIST, for screening for
XX CC antagonists or inhibitors of the interaction of MIST with cellular
XX CC signaling components. They are used in assays that detect activation
XX CC or induction of various B and T-cell-related neoplasms or cancers.
XX CC Sequences of the invention are also used in gene therapy. The present
XX CC sequence is human MIST splice variant protein from clone #12.
XX SQ Sequence 353 AA;
XX Query Match 83.9%; Score 1694; DB 23; Length 353;
XX Best Local Similarity 99.4%; Pred. No. 3.2e-143;
XX Matches 317; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 58 MEETWOSIKILPAPRIKESYADTHYFKVAMDPLDPTSTISIGOPTWNTOTRLERVD 117
```

Db 1 MEETMOSIKILPAPRIKESVADTHYFKVAMDPLPLDTRTISIGOPTWNTQTRLERYD 60
 QY 118 KPISRDVRSONIKGDASVYKNNKIPPLPPRLITLPPKKYQPLPEPSSSRPILSORHTPE 177
 Db 61 KPISRDVRSONIKGDASVYKNNKIPPLPPRLITLPPKKYQPLPEPSSSRPILSORHTPE 120
 QY 178 VQGMPSQISLRDLSEVLEAEKYPHNQKRPSTHLENTQEIPLAIISSSFTTSHSVQ 237
 Db 121 VQGMPSQISLRDLSEVLEAEKYPHNQKRPSTHLENTQEIPLAIISSSFTTSHSVQ 180
 QY 238 NRDHGGMOPCSPORCOPPASCSPHENILPYKTYTSWRPFRKSDKQVQHNEMTIGES 297
 Db 181 NRDHGGMOPCSPORCOPPASCSPHENILPYKTYTSWRPFRKSDKQVQHNEMTIGES 240
 QY 238 RQAVEAFMKENKDGSLFLVDCSTKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 357
 Db 241 RQAVEAFMKENKDGSLFLVDCSTKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 300
 QY 358 LRGEKFDSEVEDIIEHYKN 376
 Db 301 LRGEKFDSEVEDIIEHYKN 319

RESULT 5

AAB81070
 ID AAB81070 standard; Protein: 435 AA.

XX AAB81070;

DT 25-JUN-2001 (first entry)

XX Murine mast cell-specific signal transduction protein.

KW Mast cell; signal transduction; mouse; allergic disease.

OS Mus musculus.

PN JP3146204-B1.

XX 12-MAR-2001.

PD 17-SEP-1999; 99JP-0263778.

XX 17-SEP-1999; 99JP-0263778.

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI: 2001-310022/33.

DR N-PSDB; AAF86139.

PT Mast cell-specific signal-transduction molecule, useful for screening
 PT therapeutic compounds for treating allergies, is specifically expressed
 PT by mouse mast cell -

PS Claim 1; Page 7-8; 12pp; Japanese.

XX This sequence represents a murine mast cell-specific signal transduction
 CC protein. The invention includes the cDNA and protein sequences of the
 CC mast cell-specific signal transduction molecule and an expression vector
 CC containing the polynucleotide sequence. The coding sequence of the signal
 CC transduction protein can be used for screening therapeutic compounds
 CC which will be useful for treating allergic diseases.

XX Sequence 435 AA;

Query Match 56.0%; Score 1132; DB 22; Length 435;

Best Local Similarity 60.9%; Pred. No. 9.6e-93;

Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

QY 1 FONSILPNRSMRPNSTGQYQNMKRLDLMERNFAVLDGAGHSDDDYDDPELRME 60
 Db 19 FONSILPNRSMRPNSTGQYQNMKRLDLMERNFAVLDGAGHSDDDYDDPELRME 78

QY 61 TWOSIKILPAPRIKESVADTHYFKVAMDPLPLDTRTISIGOPTWNTQTRLERYD 119
 Db 79 AMPSKSIILPARIQSEVADTHYFKVAMDPLPLDTRTISIGOPTWNTQTRLERYD 138
 QY 120 ISRDVRSONIKGDASVYKNNKIPPLPPRLITLPPKKYQPLP-EPSSSRPILSORHTPE 178
 Db 139 ISRDVRSONIKGDASVYKNNKIPPLPPRLITLPPKKYQPLP-EPSSSRPILSORHTPE 198
 QY 179 VQGMPSQISLRDLSEVLEAEKYPHNQKRPSTHLENTQEIPLAIISSSFTTSHSVQ 238
 Db 199 VQGMPSQISLRDLSEVLEAEKYPHNQKRPSTHLENTQEIPLAIISSSFTTSHSVQ 258
 QY 233 RDRHGGMOPCSPORCOPPASCSPHENILPYKTYTSWRPFRKSDKQVQHNEMTIGES 298
 Db 259 RDRHGGMOPCSPORCOPPASCSPHENILPYKTYTSWRPFRKSDKQVQHNEMTIGES 316
 QY 299 RQAVEAFMKENKDGSLFLVDCSTKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 358
 Db 317 RQAVEAFMKENKDGSLFLVDCSTKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 376
 QY 359 LRGEKFDSEVEDIIEHY 374
 Db 377 LRGEKFDSEVEDIIEHY 392

RESULT 6

AAG03152
 ID AAG03152 standard; Protein: 84 AA.

XX AAG03152;

DT 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 7233.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

XX 06-SEP-2000.

PD 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

PA (GEST) GENSET.

XX Dunas Milne Edwards J, Duclert A, Giordano J;

DR WPI: 2000-500381/45.

XX N-PSDB; AAC03158.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 13; SEQ ID 7233; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream

CC regulatory sequences and to design expression and secretion vectors.
 XX Sequence 84 AA:

Query Match 17.3%; Score 349; DB 21; Length 84;
 Best Local Similarity 97.0%; Pred. No. 1.3e-23;
 Matches 64; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 FQNFSLPKRMSRPRINSATGQYORNMKPLLDMEERFAVLGAKGSHDDYDPELRME 60
 DB 19 FQNFSLPKRMSRPRINSATGQYORNMKPLLDMEERFAVLGAKGSHDDYDPELRME 78
 OY 61 TWOSIK 66
 DB 79 TWOSIK 84

RESULT 7

AAM59866 standard; Protein; 449 AA.

AC AAM59866;
 DT 20-NOV-1998 (first entry)

DE Amino acid sequence of the human BLNK-2 protein.

KW Human; BLNK-2; B cells linker protein-2; apoptosis; Grb2; PLC-gamma;
 KW SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;
 KW ras pathway; GDP; GTP; calcium pathway; antigen.

OS Homo sapiens.

PN WO9832852-A1.

PD 30-JUL-1998.

PF 23-JAN-1998; 98WO-US01394.

PR 17-MAR-1997; 97US-0819013.

PR 24-JAN-1997; 97US-0788322.

PA (UNITW) UNIV WASHINGTON.

PI Chan AC, Fu C;

DR WPI; 1998-427948/36.

DR N-PSDB; AAV41902.

PT Human B cell linker proteins - useful in the treatment of diseases

PT Involving increased or decreased apoptosis

PS Disclosure; Fig 3; 56pp; English.

This is the amino acid sequence of the human BLNK-2 (B cells linker protein-2) protein, used the treatment of diseases involving the increase and decrease of apoptosis, in the method of the invention. CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates intracellular calcium levels and Grb2 is an adapter molecule containing two SH3 domains that mediate its interaction with the CC guanine nucleotide exchange factor. Son of Sevenless (SOS) which in CC turn activates the ras pathway by facilitating the exchange of GDP for GTP on the ras molecule. Activation of both ras and calcium pathways CC are required for efficient B cell antigen receptor function. BLNK CC binding proteins can be used to identify BLNK proteins in a target sample.

XX Sequence 449 AA:

Query Match 14.3%; Score 289; DB 19; Length 449;
 Best Local Similarity 25.6%; Pred. No. 3.3e-17;
 Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;

OY 43 AKGSHDDYDPELRMEETWOSIKILPAPRIKSEYADTHYKAVMDTLPDTRTSIS 102
 DB 81 AERNADDSYEPPEVEO-----TPVPALPFRNGEY-----IDNRSS----- 118
 OY 103 GQFTWNTQTRLEKVDKPRISDVRSQNIKGDASYRKKKIPPLPPRLITLKKYQPLPEP 162
 DB 119 -----QHSPPFSKTLPSK-----PSWSEKARLTSLPALALQKPQ-VPPKP 161
 OY 163 -----ESSRP-----PLSQHRTPEVQ-----GMSQSLND 189
 DB 162 KGLLEDEADYVPEVDENDENYIHPTESSSPPEKAPVNVNSTPNSSTPASPGTASGRN 221
 OY 190 LSEVLPAEK-----VPHNORP-----ESTHLENONTOEIPLAISSSFTSNHSVQ 237
 DB 222 -SGAMETKSPPPAAPSPPLPACKKPTTLPKTTVYASQON-----ASSVCEKRP 270
 OY 238 NRDRGG--WQPCSPQRCOPASCSPHENILPY-KYT-SWRP---PFPKRSRDKVQHN 289
 DB 271 AERHRGSSHRQEAIVQSPVPFAQKQIHQKIPLPREFEGNPVDPPLPFSSTNSTISEQ 330
 OY 290 E-----WTIGESRQAYEAFMKENKDGSLVRODSTSKSEEPYLAIFYENKYYNK 342
 DB 331 EAGVLCKPWTYAGACDRKSAEALHRSNKDGSFLIRSSGHSKQPYTLVVFENKRYINP 390
 OY 343 IRFLERNOOPALGTGLRGDEKFDSDVEDIIEHYKN 376
 DB 391 VRFTEATKQYALGRKKNGEYFGSVAEIIRNHQ 424

RESULT 8

AAB93406 standard; Protein; 456 AA.

AC AAB93406;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:12602.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT full-length cDNAs -

PS Claim 8; SEQ ID 12602; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences. AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 456 AA;

Query Match 14.3%; Score 289; DB 22; Length 456;

Best Local Similarity 25.6%; Pred. No. 3.4e-17;
 Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;

OY 43 AKGSHDDYDDELMEETWOSIKILPAPRIKESYADTHYKAMDTPLDPTTSISI 102
 DB 88 AENADDSYEPPEVQOE---TRVHPALPFRAGEY-----IDNRSS--- 125
 OY 103 GQPTNNTQRLERVDKPSRDVRSNKGDSVKNKRIPLPPRLITLPKKYQPLPEP 162
 DB 126 -----QRHSPPSKTLPSK-----PSWPEKARLTSTLPALTALQKPQ-VPPKP 168
 OY 163 -----ESSRP-----PLSORHTPEVQ---GMPQSISLRD 189
 DB 169 KGLEDEADYVVPEDNDENTIHPTESSPPPEKAPMNRKSTKPSNPSPASPGTASGRN 228
 OY 190 LSEVLEAEK-----VPHNOKRP-----ESTHLENONTOEIPLAISSFTTSNHSVQ 237
 DB 229 -SGAMETKSPPPAPSPPLPRAGKKPTPLKTTVPASQON-----ASSVCEKEPIP 277
 OY 228 NRDRHGC--MPCSPQRCQPPASCSPHENILPY-KYT-SMRP-----PPPKSDKRDQVHN 289
 DB 278 AERHSGSHROAVOSVYVPPAQOKIHPKPLPRFTGEGNPTVDGPIPFSSNSTISEQ 337
 OY 290 E-----WYIGEYSROAVEAEAFMKENKDGSLVRCSTKSKEEPPYLAIFYENKYNVK 342
 DB 338 EAGVLCKRWYAGACDRKSAEELHRSNKGDSFLIRKSSGHSKQPYTLVVFENKRYNIP 397
 OY 343 IRLERNOQFALGTGLRDEKFDVSEDIIEHYKN 376
 DB 398 VRFTEATKQYALGRKKNGEEYFGSVAETIRNHQ 431

RESULT 9

AAW59865

ID AAW59865 standard; Protein; 456 AA.

AAW59865;

20-NOV-1998 (first entry)

Amino acid sequence of the human BLNK-1 protein.

Human: BLNK-1; B cells linker protein-1; apoptosis; Grb2; PLC-gamma;
 SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; Sos;
 ras pathway; GDP; GTP; calcium pathway; antigen.
 Homo sapiens.

Key Location/Qualifiers
 Misc-difference 327
 /note="encoded by AGC"

XX MO9832852-A1.
 PN 30-JUL-1998.
 PD 23-JAN-1998; 98MO-US01394.
 PF 17-MAR-1997; 97US-0819013.
 PR 24-JAN-1997; 97US-0788322.
 XX (UNITV) UNITV WASHINGTON.
 PA Chan AC, Fu C;
 PI WPI. 1998-427948/36.
 DR N-PSDB; AAW41901.
 DR Human B cell linker proteins - useful in the treatment of diseases
 PT involving increased or decreased apoptosis
 XX Claim 4; Fig 1; 56pp; English.

CC This is the amino acid sequence of the human BLNK-1 (B cells linker
 CC protein-1) protein, used the treatment of diseases involving the
 CC increase and decrease of apoptosis, in the method of the invention.
 CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
 CC intracellular calcium levels and Grb2 is an adapter molecule
 CC containing two SH3 domains that mediate its interaction with the
 CC guanine nucleotide exchange factor, Son of Sevenless (Sos) which in
 CC turn activates the ras pathway by facilitating the exchange of GDP for
 CC GTP on the ras molecule. Activation of both ras and calcium pathways
 CC are required for efficient B cell antigen receptor function. BLNK
 CC binding proteins can be used to identify BLNK proteins in a target
 CC sample.

XX Sequence 456 AA;

Query Match 14.2%; Score 286; DB 19; Length 456;

Best Local Similarity 25.6%; Pred. No. 6.2e-17;
 Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;

OY 43 AKGSHDDYDDELMEETWOSIKILPAPRIKESYADTHYKAMDTPLDPTTSISI 102
 DB 88 AENADDSYEPPEVQOE---TRVHPALPFRAGEY-----IDNRSS--- 125
 OY 103 GQPTNNTQRLERVDKPSRDVRSNKGDSVKNKRIPLPPRLITLPKKYQPLPEP 162
 DB 126 -----QRHSPPSKTLPSK-----PSWPEKARLTSTLPALTALQKPQ-VPPKP 168
 OY 163 -----ESSRP-----PLSORHTPEVQ---GMPQSISLRD 189
 DB 169 KGLEDEADYVVPEDNDENTIHPTESSPPPEKAPMNRKSTKPSNPSPASPGTASGRN 228
 OY 190 LSEVLEAEK-----VPHNOKRP-----ESTHLENONTOEIPLAISSFTTSNHSVQ 237
 DB 229 -SGAMETKSPPPAPSPPLPRAGKKPTPLKTTVPASQON-----ASSVCEKEPIP 277
 OY 228 NRDRHGC--MPCSPQRCQPPASCSPHENILPY-KYT-SMRP-----PPPKSDKRDQVHN 289
 DB 278 AERHSGSHROAVOSVYVPPAQOKIHPKPLPRFTGEGNPTVDGPIPFSSNSTISEQ 337
 OY 290 E-----WYIGEYSROAVEAEAFMKENKDGSLVRCSTKSKEEPPYLAIFYENKYNVK 342
 DB 338 EAGVLCKRWYAGACDRKSAEELHRSNKGDSFLIRKSSGHSKQPYTLVVFENKRYNIP 397
 OY 343 IRLERNOQFALGTGLRDEKFDVSEDIIEHYKN 376
 DB 398 VRFTEATKQYALGRKKNGEEYFGSVAETIRNHQ 431

RESULT 10

AAW59867

ID AAW59867 standard; Protein; 457 AA.

```

XX AC AAM59867;
XX DT 20-NOV-1998 (first entry)
XX DE Amino acid sequence of the mouse BLNK protein.
XX ID
XX KW Mouse; BLNK; B cells linker protein; apoptosis; Grb2; PLC-gamma;
XX KM SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;
XX NN ras pathway; GDP; GTP; calcium pathway; antigen.
XX OS Mus sp.
XX PN WO9832852-A1.
XX PD 30-JUL-1998.
XX PF 23-JAN-1998; 98WO-US01394.
XX PR 17-MAR-1997; 97US-0819013.
XX PR 24-JAN-1997; 97US-0788322.
XX PA (UNITW ) UNIV WASHINGTON.
XX PI Chan AC, Fu C;
XX DR WPI; 1998-427948/36.
XX DR N-PSDB; AAV41903.
XX PT Human B cell linker proteins - useful in the treatment of diseases
XX PT Involving increased or decreased apoptosis
XX PS Disclosure; Fig 5; 56pp; English.
XX XX
XX CC This is the amino acid sequence of the murine BLNK (B cells linker
XX CC protein) protein, used the treatment of diseases involving the
XX CC increase and decrease of apoptosis. In the method of the invention.
XX CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
XX CC intracellular calcium levels and Grb2 is an adapter molecule
XX CC containing two SH3 domains that mediate its interaction with the
XX CC guanine nucleotide exchange factor, Son of Sevenless (SOS) which in
XX CC turn activates the ras pathway by facilitating the exchange of GDP for
XX CC GTP on the ras molecule. Activation of both ras and calcium pathways
XX CC are required for efficient B cell antigen receptor function. BLNK
XX CC binding proteins can be used to identify BLNK proteins in a target
XX CC sample.
XX CC
XX SQ Sequence 457 AA;
Query Match 13.5%; Score 273; DB 19; Length 457;
Best Local Similarity 23.8%; Pred. No. 9.1e-16;
Matches 96; Conservative 60; Mismatches 155; Indels 92; Gaps 16;
XX 32 WERFAAVLGAAGKHS-----DDYDDPELRMEETWQSTIKILPAPRIKESV 78
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX Db 64 WSDDFSDYENPDHSDSEMYVMAEETGDSYEPPEAQ-----QTVVHPALFTTGEY 119
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX QY 79 ADTHYFVAMDPLPLDTRISISIGOPTWNTQTRLEVDKPRISDVSONIKGASVRKN 138
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX Db 120 VDNK-----SSQRHSPPFSKTLPS--KPSW-----PSAKARLASTLPADNSLOKP 162
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX QY 139 KILPPEPRPLI-----TLPKKYQLPPEP-----ESSRPLSQRHFP 176
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX Db 163 QVP-PKFKDLDEADYVVEVDNDENTIHPRSSPPAEAPVAVNRSSTKSHSP 221
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX QY 177 --EVQAMPQSLRLDLEVLAEKVPVHNOKESTHLENONTOEIPLAISSSSFTTSMH 234
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX Db 222 PGTVAGRNISGVWDSKSLPAPSPPLPRAGKKP-ATPL---KTPVP-PLPNASVCEEK 275
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX QY 235 SVQNRDHRGC--MPCSPQRCQPPASCSPHENIIPYKTKSRPPPKR----- 280
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
XX Db 276 FVPAERHRGSSRHDTQVSPVFPPTOKRVHOKVPVL-----PRPEAGSPADGPFHSF 329
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |

```

```

QY 281 -----SDRK-DVQHNENYIGETSRQAVAEAFMKENKDSFLVPOCSTKSKEEPLYAVE 333
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 330 PFNLTFADQDEGLCKPWPYAGACDRKFAEALHRSKDSFLIRSFSGHDSKOPTLVAF 389
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 334 YENKYYNKKIRFLERNOQFALGTGLRDEKFDSDVEDIIEHYNN 376
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 390 FKKRYNIPVRFLEATKOYALCKKKNGEYFSGVVEIVNSHOH 432
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
RESULT 11
ABB38147
ID ABB38147 standard; Peptide: 46 AA.
XX
AC ABB38147;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #5653 encoded by human foetal liver single exon probe.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human fetal liver -
XX PS Claim 27; SEQ ID NO 30782; 639pp + sequence listing; English.
XX
XX CC The invention relates to a single exon nucleic acid probe for
XX CC measuring human gene expression in a sample derived from human foetal
XX CC liver. The single exon nucleic acid probes may be used for predicting,
XX CC measuring and displaying gene expression in samples derived from human
XX CC fetal liver. The present invention is a peptide encoded by a single exon
XX CC nucleic acid probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_sequences.
XX
XX SQ Sequence 46 AA;
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX 311 DGSFLVDCSTKSKEEPLYAVFTEKNYVYKIRFLERNOQFALGT 356
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 1 DGSFLVDCSTKSKEEPLYAVFTEKNYVYKIRFLERNOQFALGT 46
XX | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
RESULT 12
ABB23354
ID ABB23354 standard; Protein: 46 AA.
XX

```

AC ABB23354;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #5533 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
OS Homo sapiens.
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
PS Claim 15; SEQ ID NO 25124; 530bp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA1305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular diseases, of the
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 46 AA;
XX
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 311 DGSFLVRDCSTKSKKEEPLVLAFFENKYNVKIRFLERNOQFALGT 356
DB 1 DGSFLVRDCSTKSKKEEPLVLAFFENKYNVKIRFLERNOQFALGT 46
RESULT 13
AAM58783
ID AAM58783 standard; Protein: 46 AA.
XX
AC AAM58783;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30888.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KM epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
PS Example 4; SEQ ID NO: 30888; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 46 AA;
XX
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 311 DGSFLVRDCSTKSKKEEPLVLAFFENKYNVKIRFLERNOQFALGT 356
DB 1 DGSFLVRDCSTKSKKEEPLVLAFFENKYNVKIRFLERNOQFALGT 46
RESULT 14
AAM71296
ID AAM71296 standard; Protein: 46 AA.
XX
AC AAM71296;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31602.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.

```

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 31602; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 46 AA:
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 DGSFLVRDCSTKSKKEPYVLAVFYENKYNVKIRFLERNQGFALGT 356
DB 1 DGSFLVRDCSTKSKKEPYVLAVFYENKYNVKIRFLERNQGFALGT 46

RESULT 15
AAM18973
ID AAM18973 standard; Protein; 46 AA.
XX
AC AAM18973;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #5407 encoded by probe for measuring cervical gene expression.
XX
KW Probe: human; microarray; gene expression; cervical epithelial cell;
XX
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID NO 23799; 487bp; English.
XX

```

```

CC The present invention relates to human single exon nucleic acid probes
CC (SNP: see AAI10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SNPs are derived from human HeLa cells. The SNPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 46 AA:
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 DGSFLVRDCSTKSKKEPYVLAVFYENKYNVKIRFLERNQGFALGT 356
DB 1 DGSFLVRDCSTKSKKEPYVLAVFYENKYNVKIRFLERNQGFALGT 46

```

Search completed: April 21, 2003, 12:36:25
Job time : 31.5993 secs


```

Db 88 AENADDSYPPPEOE-----TRPVHPLPFGARGEY-----IDNRSS----- 125
QY 103 GQPTNNTQRLERVDKPISRDVRSONIKGDASVRKKNKIPLPRLPLITPKKYQPLPPPP 162
Db 126 -----ORHSPPEKTLPSK-----PSWPEKARLTSTLALALAPQ--VPPRP 168
QY 163 -----ESSRP-----PLSQRTFPEVQ--GMPQSILSD 189
Db 169 KGLLEADADVVPEDNDENYIHPRESSSPPEKAPMVNRSTKPNSSSTASPGASGN 228
QY 190 LSEVLEAEK-----VPHNORKP-----ESTHLENTOTIPLAISSSTFTSNHSVQ 237
Db 229 -SGAMETSSPPAPSPPEYLRACKKPTPLKTPVASSON-----ASSVCEKPTP 277
QY 238 NRDHRGCG--MPCSPORCOPASCSPHENILPY-KYT-SWRP-----PPKRSRDRKDYGN 289
Db 278 ABRHGGSSHRQAVSPVPPPAOKOHQPIPLPRTBEGNTTVGCPDIFSSNSTISQ 337
QY 290 E-----WYIGYSRQAVEAEAFMKENKDGSLVDCSTKSKEEPLYLAVFENKVVNYK 342
Db 338 EAGVLCKRPYAGACDRKSAEALHRSNKGDSFLIRKSSHDKQPYTLVVFNNKVVNIP 397
QY 343 IRFLERNOQFALGTGRDEKDSVEDIEHYKN 376
Db 398 VREIATKOYALGRKKRNGEYGVAEIIRNHQ 431

```

RESULT 2

```

US-07-906-349A-5
; Sequence 5, Application US/07906349A
; Patent No. 3434064

```

GENERAL INFORMATION:

```

; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,349A
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 435

```

PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 07/643,237
; FILING DATE: 18-JAN-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528

```

INFORMATION FOR SEQ ID NO: 5:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-906-349A-5

```

```

Query Match 7.38; Score 147; DB 1; Length 724;
Best Local Similarity 24.78; Pred. No. 2,1e-05;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

```

```

QY 176 PEVQGMPSQISLRDISEVLEAEKVPHNQ-----RKDESTHLENTOTI- 220
Db 211 PEVQSSSEYIQL--LKKLIRSPSIPHOYWLTLQYLKHFELKLSQSSKNLNAVLSEIF 268
QY 221 -PLAISSSFTSNH-----SVONRDHRGCMQPCSPORCOPASCSPHENILPY 268
Db 269 SPMLEFRSMASSDNTENLKYVIELISTENNERQ-----PAPALPP----- 309
QY 269 KITSRRPPPK-----NSDRKDVQHNEMYIGYSRQAVEAEAFMKENKDGSLVDCSTK 322
Db 310 -----KPKKPTTVANNGMNMMNSLQNAEYWMODISREEVNEK-LRDTADGTFLVDASTK 363
QY 322 SKEEPLYLAVFENKVVNYK-IRFLERNOQFALGTGRDEKDSVEDIEHYKN 376
Db 364 MHGD-YTLTL---RKGNKKLIKIFHRQKYGFSOPL-----TFSSVELINHYRN 410

```

RESULT 3

```

US-08-167-035-2
; Sequence 2, Application US/08167035
; Patent No. 5618691

```

GENERAL INFORMATION:

```

; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: New York
; ZIP: 10036-2711

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,035
; FILING DATE: 16-DEC-1993
; CLASSIFICATION: 435

```

ATTORNEY/AGENT INFORMATION:

```

; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8664
; TELEX: 66141 PENNIE

```

INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-167-035-2

```

```

Query Match 7.38; Score 147; DB 1; Length 724;
Best Local Similarity 24.78; Pred. No. 2,1e-05;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;
QY 176 PEVQGMPSQISLRDISEVLEAEKVPHNQ-----RKDESTHLENTOTI- 220
Db 211 PEVQSSSEYIQL--LKKLIRSPSIPHOYWLTLQYLKHFELKLSQSSKNLNAVLSEIF 268
QY 221 -PLAISSSFTSNH-----SVONRDHRGCMQPCSPORCOPASCSPHENILPY 268
Db 269 SPMLEFRSMASSDNTENLKYVIELISTENNERQ-----PAPALPP----- 309

```


OY 269 KYTSMRPPPK-----RSDRKDVQHNEMWTIGEYSROAVEAFMKENKDSFLVRDOSTK 322
Db 310 -----KPKPPTVANNNGMNNMMSLQNAEMWTGDISREVEYNEK-LRDTADGTFVLVRDASTK 363
OY 323 SKEEPYLAIFYENKYNV-K-IRFLERNOQFALGTGLRGDEKFDVSVDIIIEHYKN 376
Db 364 MHGD-YTLTL--RKGNNKLKIFHRDGKYGFSDDL---TFSSVVELINHYRN 410

RESULT 4

US-08-208-887A-2
Sequence 2, Application US/08208887A
Patent No. 5677421
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Skolnick, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
KINASES AND NOVEL TARGET PROTEINS
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: 10036-2711
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/208,887A
FILING DATE: 11-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-208-887A-2

Query Match 7.3%; Score 147; DB 1; Length 724;
Best Local Similarity 24.7%; Pred. No. 2.1e-05;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

OY 176 PEVOGMSQISLRLDSEVLEAKVPHNQ-----RKPESTHLLNONTQEI- 220
Db 211 PEVOSSSEYIQL-LKKLIRSPSIPHQWLTLYLLKHFELSTSSKNLNAVLSEIF 268
OY 221 -PLAISSSFSTSNH-----SVONRDHGMQPCSPQOCPPASCPHENILPY 268
Db 269 SPMLFRSASDNTENLIKVIETLISTENNERO-----PALALP----- 309
OY 269 KYTSMRPPPK-----RSDRKDVQHNEMWTIGEYSROAVEAFMKENKDSFLVRDOSTK 322
Db 310 -----KPKPPTVANNNGMNNMMSLQNAEMWTGDISREVEYNEK-LRDTADGTFVLVRDASTK 363
OY 323 SKEEPYLAIFYENKYNV-K-IRFLERNOQFALGTGLRGDEKFDVSVDIIIEHYKN 376
Db 364 MHGD-YTLTL--RKGNNKLKIFHRDGKYGFSDDL---TFSSVVELINHYRN 410

RESULT 5

US-08-539-005-2
Sequence 2, Application US/08539005
Patent No. 5858686
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Skolnick, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
KINASES AND NOVEL TARGET PROTEINS
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: 10036-2711
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,005
FILING DATE: 4-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/167,035
FILING DATE: 16-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-539-005-2

Query Match 7.3%; Score 147; DB 2; Length 724;
Best Local Similarity 24.7%; Pred. No. 2.1e-05;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

OY 176 PEVOGMSQISLRLDSEVLEAKVPHNQ-----RKPESTHLLNONTQEI- 220
Db 211 PEVOSSSEYIQL-LKKLIRSPSIPHQWLTLYLLKHFELSTSSKNLNAVLSEIF 268
OY 221 -PLAISSSFSTSNH-----SVONRDHGMQPCSPQOCPPASCPHENILPY 268
Db 269 SPMLFRSASDNTENLIKVIETLISTENNERO-----PALALP----- 309
OY 269 KYTSMRPPPK-----RSDRKDVQHNEMWTIGEYSROAVEAFMKENKDSFLVRDOSTK 322
Db 310 -----KPKPPTVANNNGMNNMMSLQNAEMWTGDISREVEYNEK-LRDTADGTFVLVRDASTK 363
OY 323 SKEEPYLAIFYENKYNV-K-IRFLERNOQFALGTGLRGDEKFDVSVDIIIEHYKN 376
Db 364 MHGD-YTLTL--RKGNNKLKIFHRDGKYGFSDDL---TFSSVVELINHYRN 410

RESULT 6
US-09-280-598-5

Sequence 5, Application US/09280598
Patent No. 6391584
GENERAL INFORMATION:
APPLICANT: Schlusser, Joseph
APPLICANT: Scholnik, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
KINASES AND NOVEL TARGET PROTEINS
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/280,598
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/252,820
FILING DATE: 02-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-067
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-280-598-5

Query Match 7.3%, Score 146.5; DB 4; Length 724;
Best Local Similarity 24.7%, Pred. No. 2.3e-05;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

QY 176 PEVOGMPQSISRDLSEVLEAEKVPHNQ-----RKPESTHLLNONTQET- 220
| | | | | : : : : : | |
DB 211 PEVOSSSEYIQL-LKKLIRSPSIHQYWLTLQYLKHFPLKSQTSKNLNLARVLSLTF 268
| | | | | : : : : : | |
QY 221 -PLAIISSSFTTSNH-----SVQNRDRHGMOPCSPORCOPASPSPHENILPY 268
| : : : : : | : : : : : | |
DB 269 SPMLFRFSAASDNTENLIKVEILISTEWMNRQ-----PAPALRP----- 309
| : : : : : | : : : : : | |
QY 269 KYTSRRPFRP-----KSDRKDVQHNENYIGEYSRQAVEAFMKENKDGSLVADCSK 322
| : : : : : | : : : : : | |
DB 310 -----KPKRPPTTVANNNGNNNNNSLQNAEMWGDISRREYNEK-LRDTADGFLVADASTK 363
| : : : : : | : : : : : | |
QY 323 SKEEPYVAVFENKVVNVK-IRFLERNQOFALGTGLRDEKFDSEVEDIIEHYKN 376
| : : : : : | : : : : : | |
DB 364 MHGD-YTILTL---RKGGNNKLIKIRHROGKIGFSDPL-----TFSSVELINHYRN 410
| : : : : : | : : : : : | |

RESULT 7
US-08-390-874C-12
Sequence 12, Application US/08390874C
GENERAL INFORMATION:
APPLICANT: Klippel, Anke

APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: A Constitutively Active
Phosphatidylinositol 3-kinase and Uses Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,874C
FILING DATE: 17-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 02307K-057000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 722 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-874C-12

Query Match 7.2%, Score 144.5; DB 3; Length 722;
Best Local Similarity 21.4%, Pred. No. 3.6e-05;
Matches 75; Conservative 40; Mismatches 95; Indels 141; Gaps 17;

QY 141 PLPPRPPLITLPPKYYQPLPEPESR-----PPLSQRTPEVOGMPQSISR 188
| | | | | : | | | | : | | | |
DB 85 PTPKPRP---PR---PLPVAGSSKTEADTEQALPLPDLAEQFAPPDV-APLLIKL 136
| | | | | : | | | | : | | | |
QY 189 DLSEV-LEAEVPHNQKPESTHLE-----NONTQETPLAIISSSFTTSNHSVONR 239
| | | | | : | | | | : | | | |
DB 137 EAIEKKGLCESTLYRTQSSNPALRQLDDCAASVDELMIDVHLADAFKRYLADLPN- 195
| | | | | : | | | | : | | | |
QY 240 DHRGMOPCSP-----QRCQPPASC-----SPHENILPYKY----- 270
| | | | | : | | | | : | | | |
DB 196 -----FVIRPVAVYNNEMSLAQELQSPEDCIQLKLRPLNPTRHQCMVTLQYLKHF 248
| | | | | : | | | | : | | | |
QY 271 -----TSW----- 273
| | | | | : | | | | : | | | |
DB 249 KLSQASSKNLNLARVLSLTFSPVLFRRPAASDNTENLIKAEILISTEWMNRQAPALP 308
| : : : : : | : : : : : | |
QY 274 -RPPRP-----KSDRKDVQHNENYIGEYSRQAVEAFMKENKDGSLVADCSKSE 326
| : : : : : | : : : : : | |
DB 309 PKRPPTTVANNNGNNNNNSLQNAEMWGDISRREYNEK-LRDTADGFLVADASTKMG 366
| : : : : : | : : : : : | |
QY 327 PYVAVFENKVVNVK-IRFLERNQOFALGTGLRDEKFDSEVEDIIEHYKN 376
| : : : : : | : : : : : | |
DB 367 -YTILTL---RKGGNNKLIKIRHROGKIGFSDPL-----TFNSVELINHYRN 409
| : : : : : | : : : : : | |

RESULT 8
US-09-265-772-12
Sequence 12, Application US/09265772
Patent No. 6300111
GENERAL INFORMATION:
APPLICANT: Klippel, Anke
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: A Constitutively Active

TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,772
FILING DATE: 10-MAR-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,874
FILING DATE: 17-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 02307K-057010US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 722 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-265-772-12

Query Match 7.2%; Score 144.5; DB 4; Length 722;
Best Local Similarity 21.4%; Pred. No. 3.6e-05;
Matches 75; Conservative 40; Mismatches 95; Indels 141; Gaps 17;
QY 141 PLPPPPPLTLPRKKYOLPPEPSSR-----PPLSGRHFFPEYQGMPSQISLR 188
DB 85 PPKPKPP-----PR---PLPVAFGSSKTEADTEGOALPLPLABQFAPDV-APPLIKL 136
QY 189 DLSEV--LEAEKYPHNRPESTHLE-----NNTQEIPLAIISSSTTNSHVNQR 239
DB 137 EAEKGLGECSTLYRTQSSSNPAELRQLDDCAASVDLEMDIVHADAKRYLADLPN- 195
QY 240 DHRGMQPCSP-----ORCOPASC-----SPHENILPYKY----- 270
DB 196 -----PIIPVAVYVEMMSLAQELDSPEDCIOLKKLIRLPNIPRQCHWTLOYLLKHF 248
QY 271 -----TSW----- 273
DB 249 KLSQASSKNILNARVLEIFSPVLFRRPAAASDNTHELKAIKIELISTEMNERQAPALP 308
QY 274 -RRPF-----KSRDKDYQHNEMWIGESRQAVEAEFAKKEKDSFLYRDCSTSKKE 326
DB 309 PPKPKPTTVAANSMNNMMSLQDAEWGDISREYNK--LRDTADGTFTFLRDASTYKMGD 366
QY 327 PYLAIVFENKYYNVK-IRFLERNQOFALGTGLRGDEKFDVEDIEHYKN 376
DB 367 -YTLTL--RKGGNKLKIFHRDQKYGFSPL-----TNSVVELLNHYRN 409

RESULT 9
US-08-391-615-2
Sequence 2, Application US/08391615
Patent No. 5550054
GENERAL INFORMATION:
APPLICANT: Witte, Owen
APPLICANT: Tsukada, Satoshi

APPLICANT: Saffran, Douglas
APPLICANT: Rawlings, David
TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHNACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,615
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/006,449
FILING DATE: 21-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Berttram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-391-615-2

Query Match 7.1%; Score 143; DB 1; Length 659;
Best Local Similarity 23.4%; Pred. No. 4.4e-05;
Matches 60; Conservative 27; Mismatches 71; Indels 98; Gaps 11;
QY 125 RSNINIGDASVKNKPIPLPPRPLTLPRKKYOLPPEPSSRPLSGRHFFPEYQGMPSQ 184
DB 171 RNSGLKPGSSHRKTRKPLPTPEDDILK--PLPEPTA-----AP 210
QY 185 ISLRDSEVLEA-EKYPHNO-----RRPESTHLENONFOEIPLAIISSSTTNSHVNQR 238
DB 211 ISTTELKRYVALDYMPMANNNDQLRKGEYFLESN--LW-----WRA 254
QY 239 RDHRGMQPCSPQRCOPASCSPHENILPYKYTSWRPPEKRSRDKRDYQHNEMWIGESYR 298
DB 255 RDKNG-----QEGYIPSNYIT-----DAEDSIEMEYNSHMTR 288
QY 299 QAVEEAPMKENKDGSLYRDCSTKSKKEPYLAIVFENKYYNKKIRPLENNOQFALGTGL 358
DB 289 SQAEOELKQKGGKGGIVRSKAGK-----YTVSV-----FAKSTG- 325
QY 359 RGDEKFDVEDIEHY 374
DB 326 -----EPQGVIRHY 334

RESULT 10
US-08-426-509A-8
Sequence 8, Application US/08426509A
Patent No. 6326469
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Gishitsky, Mikhail
APPLICANT: Sures, Irman G.

```

; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
; TITLE OF INVENTION: TYROSINE KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York,
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,509A
; FILING DATE: 21-Apr-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/232,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-0074-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: No. 6326469e
; US-08-426-509A-8

```

Query Match 6.9%; Score 140; DB 4; Length 659;

Best Local Similarity 22.7%; Pred. No. 8.5e-05; Matches 58; Conservative 29; Mismatches 71; Indels 98; Gaps 11;

```

QY 125 RSONIKGDASYRNKKIPLPPRPRLITLPPKYYQPLPPEPSSRPPLSQRHFRPEYQGMPSQ 184
Db 171 RNSGLKGGSSHRKTKKPLPRPEEDQILK--PLPPEPAA-----Ap 210
QY 185 ISLSDLSEVLEA-EKVPHNQ-----RKPESTHLELNONTOEIPLAISSSFTTSNHSYON 238
Db 211 VSTSELKKVVALYDYMPMNANDLQLRKGDYEFILLESN--LPW-----WRA 254
QY 239 RDHNGMOPCSPQPCOPPASCSPHENILPYKYTSWRRPFRKSRDKDQVHNEMWIGEYSR 298
Db 255 RDKNG-----OEGYIPSNYVI-----EAEDSTLEMWEMSKHMT 288
QY 299 QAVEAEPKKNKDGSLVRDCSTKSKEPPYLAIFYENKYNVVKIRLERNOQFALGTGL 358
Db 289 SQAEQLLKQEGKEGGFIVRDSKAGK-----YTVSV-----FAKSTG- 325
QY 359 RGDEKFDSEVDIIIEHY 374
Db 326 -----DPOGVIRHY 334

```

RESULT 11
PCT-US95-05008-8
Sequence 8, Application PC/TUS9505008

GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.

```

; APPLICANT: Hofgarten Str. 2
; APPLICANT: Munchen 80539
; APPLICANT: Germany
; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05008
; FILING DATE: 24-Apr-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/232,545
; FILING DATE: 22-Apr-1994
; TELECOMMUNICATION INFORMATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-074
; TELEPHONE: (212)790-9090
; TELEFAX: (212)869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-05008-8

```

Query Match 6.9%; Score 140; DB 5; Length 659;

Best Local Similarity 22.7%; Pred. No. 8.5e-05; Matches 58; Conservative 29; Mismatches 71; Indels 98; Gaps 11;

```

QY 125 RSONIKGDASYRNKKIPLPPRPRLITLPPKYYQPLPPEPSSRPPLSQRHFRPEYQGMPSQ 184
Db 171 RNSGLKGGSSHRKTKKPLPRPEEDQILK--PLPPEPAA-----Ap 210
QY 185 ISLSDLSEVLEA-EKVPHNQ-----RKPESTHLELNONTOEIPLAISSSFTTSNHSYON 238
Db 211 VSTSELKKVVALYDYMPMNANDLQLRKGDYEFILLESN--LPW-----WRA 254
QY 239 RDHNGMOPCSPQPCOPPASCSPHENILPYKYTSWRRPFRKSRDKDQVHNEMWIGEYSR 298
Db 255 RDKNG-----OEGYIPSNYVI-----EAEDSTLEMWEMSKHMT 288
QY 299 QAVEAEPKKNKDGSLVRDCSTKSKEPPYLAIFYENKYNVVKIRLERNOQFALGTGL 358
Db 289 SQAEQLLKQEGKEGGFIVRDSKAGK-----YTVSV-----FAKSTG- 325
QY 359 RGDEKFDSEVDIIIEHY 374
Db 326 -----DPOGVIRHY 334

```

RESULT 12
US-08-391-615-6
Sequence 6, Application US/08391615
Patent No. 5550054
GENERAL INFORMATION:

APPLICANT: Witte, Owen
APPLICANT: Tsukada, Satoshi
APPLICANT: Saffran, Douglas
APPLICANT: Rawlings, David
TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,615
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/006,449
FILING DATE: 21-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Berttram I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 RHT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-615-6

Query Match 6.8%; Score 137; DB 1; Length 442;
Best Local Similarity 24.9%; Pred. No. 9.1e-05;
Matches 54; Conservative 27; Mismatches 61; Indels 75; Gaps 9;
QY 125 RSQNIKGDASVRKKNKIPPLPPRLITLPPKYOPLPEPSSRPPLSORHTEPEV---OGM 181
DB 171 RNSGLKPGSSHRKTKKPLPPTPEEDQILK--PLPPEPAA--PVSTSELKKEVALDYM 226
QY 182 P---SQSLRDLSE--VLEAKVPHNQKRPSTHLENONOEPLAISSSFTTSNSV 236
DB 227 PMANNDQLRKGDEFFLEESNLLPWWRAR-----DKNOE-----GYLPSNDVT 270
QY 237 QNRDHRCGMQPCSPQRCPPASCSPHENILPYKYSWPRPPKRSRDKVOHNEWYIDEY 296
DB 271 EAED-----SIETMYTSKRM 286
QY 297 SRQAVEAFMKENKDGSLVRCSTKRSKEEPLYAVF 333
DB 287 TRSOAEOLLKQEGKEGFIVDDSSKAAR---YTLVSF 320

RESULT 13
US-08-202-389-28
Sequence 28, Application US/08202389
Patent No. 5536636
GENERAL INFORMATION:
APPLICANT: Freeman Jr., Robert M.
APPLICANT: Plutzky, Jorge
APPLICANT: Neel, Benjamin G.
APPLICANT: Rosenberg, Robert D.

TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,389
FILING DATE: 28-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,926
FILING DATE: 01-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/829,141
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/721,112
FILING DATE: 26-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: BIH92-05WA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-202-389-28

Query Match 6.4%; Score 130; DB 1; Length 92;
Best Local Similarity 34.1%; Pred. No. 4.2e-05;
Matches 29; Conservative 26; Mismatches 22; Indels 8; Gaps 3;
QY 291 WYIGYSQAVEAFMKENKDGSLVRCSTKRSKEEPLYAVFENKYNNKIRFLERNO 350
DB 1 WFGHLSKREKLEKILLKNGKSGFLVRE--SQSKRGDFVLSVRDDKVTYHWR--QDK 56
QY 351 QFALGTGLRGDEKEDSVEDIEIHYK 375
DB 57 KYDVG---GGSEFGTSELIDHYK 77

RESULT 14
US-08-729-416C-1
Sequence 1, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-1

```

```

Query Match          6.4%; Score 129.5; DB 3; Length 474;
Best Local Similarity 24.7%; Pred. No. 0.00052;
Matches 72; Conservative 37; Mismatches 112; Indels 71; Gaps 16;

```

```

QY 112 RLRRVVKPIRSDVRSQINIGDASVRK--NKIRLPPRPRLITLTKKQYPLRPPESS----165
DB 201 RMOSLDEPWTEE-----EDGSDHPYNSIPSKMPGGFDTLRKPRHAPDQAQFAG 234
QY 166 --RPPLSQRH---TFPE-VQGMPSQISLRDLSEVLEAE-KVPHNQKPESTHLENNQ 218
DB 255 KEQTYVQGRRLGDTFEGEDMQOTPLRQSSDIYSTPEGKLHVATGEP--TV-----NTQ 308
QY 219 ET-----PLAIISSSFTTSMHVSQVQNDHRCGM--QPCSP-----ORCPAPASCSPHEN 264
DB 309 QIPQAMPAAVSSAESSPRKDLFDMKPFEDALKNQPLGPIVLSKAASVEICISPV-----362
QY 265 ILPYKXTSMRPPPKRSDRQVQHNEMVYIGEVSQVQAEVEAFMKNENKDGSLVYDCSTKSK 324
DB 363 -----PRAPDAKMLEELQAEIETWYQEGMSRKEAEGIL---EKDGDVLR---KST 405
QY 325 EEP--YVLAVFYENKYVNVKIRFLERNOQFALGTGLRGDEKFDSDIEHY 374
DB 406 TNGGSFVLGMHNGQAKHLLVDP-----GTIRTKDVFDSISHLINHH 450

```

RESULT 15

US-08-729-416C-7

Sequence 7, Application US/08729416C

Patent No. 6013767

GENERAL INFORMATION:

APPLICANT: NAKAMURA, TAKESHI

TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE

TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.

STREET: 1100 NEW YORK AVENUE, N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/729,416C

FILING DATE: 11-OCT-1996

```

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-7

```

```

Query Match          6.4%; Score 129.5; DB 3; Length 594;
Best Local Similarity 24.7%; Pred. No. 0.00072;
Matches 72; Conservative 37; Mismatches 112; Indels 71; Gaps 16;

```

```

QY 112 RLRRVVKPIRSDVRSQINIGDASVRK--NKIRLPPRPRLITLTKKQYPLRPPESS----165
DB 321 RMOSLDEPWTEE-----EDGSDHPYNSIPSKMPGGFDTLRKPRHAPDQAQFAG 374
QY 166 --RPPLSQRH---TFPE-VQGMPSQISLRDLSEVLEAE-KVPHNQKPESTHLENNQ 218
DB 375 KEQTYVQGRRLGDTFEGEDMQOTPLRQSSDIYSTPEGKLHVATGEP--TV-----NTQ 428
QY 219 ET-----PLAIISSSFTTSMHVSQVQNDHRCGM--QPCSP-----ORCPAPASCSPHEN 264
DB 429 QIPQAMPAAVSSAESSPRKDLFDMKPFEDALKNQPLGPIVLSKAASVEICISPV-----482
QY 265 ILPYKXTSMRPPPKRSDRQVQHNEMVYIGEVSQVQAEVEAFMKNENKDGSLVYDCSTKSK 324
DB 483 -----PRAPDAKMLEELQAEIETWYQEGMSRKEAEGIL---EKDGDVLR---KST 525
QY 325 EEP--YVLAVFYENKYVNVKIRFLERNOQFALGTGLRGDEKFDSDIEHY 374
DB 526 TNGGSFVLGMHNGQAKHLLVDP-----GTIRTKDVFDSISHLINHH 570

```

Search completed: April 21, 2003, 12:39:08
Job time: 15.5906 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:36:32 ; Search time 13.4451 Seconds
(without alignments)
2114.581 Million cell updates/sec

Title: US-09-856-061-4
Perfect score: 2020
Sequence: 1 FQNFSLPKRNSWPRINSATG.....GLRGEKEDVEDIEHYKN 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2009	99.5	428	9	US-09-966-955A-4
2	2009	99.5	443	9	US-09-966-955A-2
3	1694	83.9	353	9	US-09-966-955A-6
4	238	11.8	46	10	US-09-864-761-38652
5	214	10.6	40	10	US-09-864-761-44023
6	149	7.4	688	9	US-10-081-9808-1
7	147	7.3	724	10	US-09-962-929-2
8	146.5	7.2	724	10	US-09-962-929-4
9	144.5	7.2	945	8	US-08-965-272-2
10	144.5	7.2	945	12	US-10-121-882-2
11	143	7.1	659	9	US-10-045-202-4
12	140	6.9	659	9	US-09-977-260-8
13	140	6.9	659	9	US-10-045-202-2
14	140	6.9	659	9	US-09-977-261-8
15	140	6.9	659	10	US-09-977-261-8
16	126	6.2	597	9	US-10-038-010-22
17	125.5	6.2	533	10	US-09-908-805B-63
18	125	6.2	593	10	US-09-920-021A-3

20	120	5.9	939	10	US-09-226-248B-29	Sequence 29, Appl
21	120	5.9	939	10	US-09-801-368-380	Sequence 380, App
22	119	5.9	620	9	US-09-977-260-9	Sequence 9, Appl
23	119	5.9	620	9	US-09-977-261-9	Sequence 9, Appl
24	119	5.9	620	10	US-09-977-269-9	Sequence 9, Appl
25	115	5.7	675	9	US-10-186-399-3	Sequence 3, Appl
26	115	5.7	675	9	US-09-977-260-4	Sequence 4, Appl
27	115	5.7	675	9	US-09-977-261-4	Sequence 4, Appl
28	115	5.7	675	10	US-09-977-269-4	Sequence 4, Appl
29	112.5	5.6	552	10	US-09-880-192-50	Sequence 50, Appl
30	111	5.5	542	10	US-09-908-805B-44	Sequence 44, Appl
31	111	5.5	655	10	US-09-205-658-57	Sequence 57, Appl
32	111	5.5	655	10	US-09-844-353A-57	Sequence 57, Appl
33	111	5.5	944	10	US-09-964-238-2	Sequence 2, Appl
34	110	5.4	1341	9	US-10-076-622-565	Sequence 565, App
35	110	5.4	1341	12	US-10-007-805-565	Sequence 565, App
36	110	5.4	1349	9	US-10-076-622-573	Sequence 573, App
37	110	5.4	1349	12	US-10-007-805-573	Sequence 573, App
38	109.5	5.4	4019	9	US-09-854-133-425	Sequence 425, App
39	109.5	5.4	4019	10	US-09-738-973-425	Sequence 425, App
40	107.5	5.3	365	9	US-10-029-180-62	Sequence 62, Appl
41	107.5	5.3	1191	10	US-09-921-099-2	Sequence 2, Appl
42	107.5	5.3	1191	10	US-09-921-099-4	Sequence 4, Appl
43	107.5	5.3	1325	10	US-09-864-761-35612	Sequence 35612, A
44	106	5.2	673	10	US-09-727-384-4	Sequence 4, Appl
45	105.5	5.2	479	10	US-09-881-752A-312	Sequence 312, App

ALIGNMENTS

RESULT 1
US-09-966-955A-4
; Sequence 4, Application US/0966955A
; Patent No. US2002015563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gene S.
; APPLICANT: Kanter, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-Length Human
; TITLE OF INVENTION: Cink-related Gene, MIST (Mast Cell Immunoreceptor)
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113051
; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #7, AMINO ACID
; OTHER INFORMATION: SEQUENCE
US-09-966-955A-4

Query Match 99.5% Score 2009; DB 9; Length 428;
Best Local Similarity 99.5% Pred. No. 8.7e-142;
Matches 374; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQNFSLPKRNSWPRINSATGQYORMNKPLDWERNEFAVLGAKGSHDDDDYDDELPMEE 60
Db 19 FQNFSLPKRNSWPRINSATGQYORMNKPLDWERNEFAVLGAKGSHDDDDYDDELPMEE 78
QY 61 TWOSIKILPAPRIKESYADTHYFKVAMDPLPLDTRTSISISQPTWNTOTRLERVDKPI 120
Db 79 TWOSIKILPAPRIKESYADTHYFKVAMDPLPLDTRTSISISQPTWNTOTRLERVDKPI 138

QY 121 SRDVSQNIKGDASVRKNNKIPPLPPRPLITLTKKYOTPLPPEPSSRPLSQRHTPEVQG 180
DB 129 SKDVASQNIKGDASVRKNNKIPPLPPRPLITLTKKYOTPLPPEPSSRPLSQRHTPEVQG 198
QY 161 MFSQISLRLDLSEVLEAEKVPYVPHNQRKPESTHLENNONTOEIPLAISSTTSNHSVQND 240
DB 199 MFSQISLRLDLSEVLEAEKVPYVPHNQRKPESTHLENNONTOEIPLAISSTTSNHSVQND 258
QY 241 HRGMOQPCSPORCOPASCSPHENILPYKYSWRPFPKRSRDKDQVQHNEMWIGEYSRQA 300
DB 259 HRGMOQPCSPORCOPASCSPHENILPYKYSWRPFPKRSRDKDQVQHNEMWIGEYSRQA 318
QY 301 VEEAFMKENKDGSLVLRDCKSTKSKEEPPYLAIFYENKYNVAKIRFLERNQOFALGTGLRG 360
DB 319 VEEAFMKENKDGSLVLRDCKSTKSKEEPPYLAIFYENKYNVAKIRFLERNQOFALGTGLRG 378
QY 361 DEKFDSEVDIIIEHYKN 376
DB 379 DEKFDSEVDIIIEHYKN 394

RESULT 2

US-09-966-955A-2
Sequence 2, Application US/09966955A
Patent No. US20020155563A1
GENERAL INFORMATION:
APPLICANT: Perez-Villar, Juan J.
APPLICANT: Chang, Han
APPLICANT: Yang, Wen-Pin
APPLICANT: Wu, Yuli
APPLICANT: Whitley, Gena S.
APPLICANT: Kanner, Steven B.
TITLE OF INVENTION: Identification and Cloning of a Full-length Human
TITLE OF INVENTION: Clonk-related Gene, MIST (Mast Cell Immunoceptor)
TITLE OF INVENTION: Signal Transducer)
FILE REFERENCE: 3053-4113051
CURRENT APPLICATION NUMBER: US/09/966, 955A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/237030
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 443
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
OTHER INFORMATION: HUMAN FULL-LENGTH MIST cDNA CLONE #8, TRANSLATED
OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-2

Query Match 99.5%; Score 2009; DB 9; Length 443;
Best Local Similarity 99.5%; Pred. No. 9.1e-142;

Matches 374; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FONSLPKNRSRPRINSATGQYORANKPLDWERNFVAVLDOGAKSHDDYDDPELRME 60
DB 34 FONSLPKNRSRPRINSATGQYORANKPLDWERNFVAVLDOGAKSHDDYDDPELRME 93
QY 61 TWQSIKILPARIKESVADTHYFKVAMDPLPLDTRTSISIGOPTWNTQTRLEVRDPI 120
DB 94 TWQSIKILPARIKESVADTHYFKVAMDPLPLDTRTSISIGOPTWNTQTRLEVRDPI 153
QY 121 SRDVSQNIKGDASVRKNNKIPPLPPRPLITLTKKYOTPLPPEPSSRPLSQRHTPEVQG 180
DB 154 SKDVASQNIKGDASVRKNNKIPPLPPRPLITLTKKYOTPLPPEPSSRPLSQRHTPEVQG 213
QY 181 MFSQISLRLDLSEVLEAEKVPYVPHNQRKPESTHLENNONTOEIPLAISSTTSNHSVQND 240
DB 214 MFSQISLRLDLSEVLEAEKVPYVPHNQRKPESTHLENNONTOEIPLAISSTTSNHSVQND 273
QY 241 HRGMOQPCSPORCOPASCSPHENILPYKYSWRPFPKRSRDKDQVQHNEMWIGEYSRQA 300

DB 274 HRGMOQPCSPORCOPASCSPHENILPYKYSWRPFPKRSRDKDQVQHNEMWIGEYSRQA 333
QY 301 VEEAFMKENKDGSLVLRDCKSTKSKEEPPYLAIFYENKYNVAKIRFLERNQOFALGTGLRG 360
DB 334 VEEAFMKENKDGSLVLRDCKSTKSKEEPPYLAIFYENKYNVAKIRFLERNQOFALGTGLRG 393
QY 361 DEKFDSEVDIIIEHYKN 376
DB 394 DEKFDSEVDIIIEHYKN 409

RESULT 3

US-09-966-955A-6
Sequence 6, Application US/09966955A
Patent No. US20020155563A1
GENERAL INFORMATION:
APPLICANT: Perez-Villar, Juan J.
APPLICANT: Chang, Han
APPLICANT: Yang, Wen-Pin
APPLICANT: Wu, Yuli
APPLICANT: Whitley, Gena S.
APPLICANT: Kanner, Steven B.
TITLE OF INVENTION: Identification and Cloning of a Full-length Human
TITLE OF INVENTION: Clonk-related Gene, MIST (Mast Cell Immunoceptor)
TITLE OF INVENTION: Signal Transducer)
FILE REFERENCE: 3053-4113051
CURRENT APPLICATION NUMBER: US/09/966, 955A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/237030
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 353
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #12, TRANSLATED
OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-6

Query Match 83.9%; Score 1694; DB 9; Length 353;
Best Local Similarity 99.4%; Pred. No. 1.8e-118;

Matches 317; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 58 MEETQSIKILPARIKESVADTHYFKVAMDPLPLDTRTSISIGOPTWNTQTRLEVRD 117
DB 1 MEETQSIKILPARIKESVADTHYFKVAMDPLPLDTRTSISIGOPTWNTQTRLEVRD 60
QY 118 KRISDVASQNIKGDASVRKNNKIPPLPPRPLITLTKKYOTPLPPEPSSRPLSQRHTPE 177
DB 61 KRISDVASQNIKGDASVRKNNKIPPLPPRPLITLTKKYOTPLPPEPSSRPLSQRHTPE 120
QY 178 VQGMFSQISLRLDLSEVLEAEKVPYVPHNQRKPESTHLENNONTOEIPLAISSTTSNHSVQ 237
DB 121 VQGMFSQISLRLDLSEVLEAEKVPYVPHNQRKPESTHLENNONTOEIPLAISSTTSNHSVQ 180
QY 238 NRDHGMOQPCSPORCOPASCSPHENILPYKYSWRPFPKRSRDKDQVQHNEMWIGEYS 297
DB 181 NRDHGMOQPCSPORCOPASCSPHENILPYKYSWRPFPKRSRDKDQVQHNEMWIGEYS 240
QY 298 ROAVEAFMKENKDGSLVLRDCKSTKSKEEPPYLAIFYENKYNVAKIRFLERNQOFALGTG 357
DB 241 ROAVEAFMKENKDGSLVLRDCKSTKSKEEPPYLAIFYENKYNVAKIRFLERNQOFALGTG 300
QY 358 LRGDSEVDIIIEHYKN 376
DB 301 LRGDSEVDIIIEHYKN 319

RESULT 4

US-09-864-761-38652
Sequence 38652, Application US/09864761


```
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 38652
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005599.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
OTHER INFORMATION: SWISSPROT HIT: Q60787, EVALUATION 9.00e-10
US-09-864-761-38652

Query Match 11.8%; Score 238; DB 10; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 5
US-09-864-761-44023
Sequence 44023, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 44023
LENGTH: 40
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005599.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.72
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.68
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.47
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
OTHER INFORMATION: SWISSPROT HIT: P25502, EVALUATION 3.70e+00
US-09-864-761-44023

Query Match 10.6%; Score 214; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Db 196 -----PVIPAVNEMMSLAOELSPEDCIOLKLLRLPNIPHCWTLQYLKHF 248
QY 271 -----TSW----- 273
Db 249 KLSQASSKULNARVSEIFSPVLFPPAASSNTEHLIAEILLISTENNEROPAPALP 308
QY 274 -RPPFP-----RSDRKDVONHEWYIGEXSROAVEAFENKENDGSLVRDCSTKSEE 326
Db 309 KPPKFTYVANNMNNMNSIQDAEWYWGDISREVEVEK-LRDTADGTFVLVRDASTKMHGD 367
QY 327 PYLAVFYENKYNVK-IRFLERNOQFALGTGRGDEKFPDVEDIIEHYKN 376
Db 368 -YTLT---PRKGGNNKLIKIFHRDGKYGFSDDL-----TFNSVVELINHYRN 410

RESULT 9
US-09-962-929-4
Sequence 4, Application US/09962929
Patent No. US20020115058A1
GENERAL INFORMATION:
APPLICANT: Pedersen, Finn S.
APPLICANT: Soerensen, Annette B.
APPLICANT: Nielsen, Anne A.
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Diseases Associated with A
FILE REFERENCE: A-70004/RMS/DCF
CURRENT APPLICATION NUMBER: US/09/962,929
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 09/668,644
SOFTWARE: Patent version 3.1
SEQ ID NO 4
LENGTH: 724
TYPE: PRT
ORGANISM: Homo sapiens
US-09-962-929-4

Query Match
Best local Similarity 7.3%; Score 146.5; DB 10; Length 724;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

QY 176 PEVOGHPQISLRDLSEVLEAEKVPHNQ-----RKPESTHLENOQTQEI- 220
Db 211 PEVOSSSEYIQL--IKKLIRSPISPHQYMLTQYLKHFKLSQTSKLNARVSEIF 268
QY 221 -PLAIISSSTFTSNH-----SVQNRDRHGGMQPCSPQRCOPPASCSEHENTLPY 268
Db 269 SPMLFRFSASSDNTENLIRVIELISTENNERO-----PAPALP----- 309
QY 269 KYTSWRPPFPK-----RSDRKDVONHEWYIGEXSROAVEAFENKENDGSLVRDCSTK 322
Db 310 -----KPPKFTYVANNMNNMNSIQDAEWYWGDISREVEVEK-LRDTADGTFVLVRDASTK 363
QY 323 KEEPYLAVFYENKYNVK-IRFLERNOQFALGTGRGDEKFPDVEDIIEHYKN 376
Db 364 MHGD-YTLT---RKGGNNKLIKIFHRDGKYGFSDDL-----TFNSVVELINHYRN 410

RESULT 10
US-08-965-272-2
Sequence 2, Application US/08965272
Patent No. US20020019514A1
GENERAL INFORMATION:
APPLICANT: Glimcher, Laurie H. et al.
TITLE OF INVENTION: CITRA-INTERACTING PROTEINS AND METHODS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts

COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,272
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: 41,106
REFERENCE/DOCKET NUMBER: HUI-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 945 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-965-272-2

Query Match
Best local Similarity 7.2%; Score 144.5; DB 8; Length 945;
Matches 64; Conservative 22; Mismatches 80; Indels 81; Gaps 12;

QY 53 DPELRMETWQSIK--ILPAPRIKESYADTHYFKVAMDTPLD---TFTSISIGOPTW 107
Db 5 DESIRKEEQOQHAGVAPQPLKEP-----FASLOSPPFDPAPTTAAVAVATT 56
QY 108 NFOIRLEKVDKPSRDVRSQNIKGDASVRKNKIP---LPPRPPLITLPKKYOPLPPEPES 164
Db 57 TTTTITTT-----TAQDEKKRPPALPPPPPLAKPPPSQOPPPP-- 97
QY 165 SREPLSGRHTPEVOGHPQIS-LRDLSEVLEAEK-----VPHNRKP- 206
Db 98 -PP-----PSPASLKLASVLEGQKCYRGTAAGAVSTRPGPLPTTQYSPG 142
QY 207 -----ESTHLENOQTQEIPLAISSTFTSNHVSQNRDRHGGMQ-----PCSPQRCQ 254
Db 143 PPSGATALPPTSAAPSAGSPPSASSSOSTSG-GPMARERRAGEEVPVGPMTPTOPP 201
QY 255 PPASCSP 261
Db 202 PPLSLPP 208

RESULT 11
US-10-121-882-2
Sequence 2, Application US/10121882
Patent No. US20020146806A1
GENERAL INFORMATION:
APPLICANT: Glimcher, Laurie H. et al.
TITLE OF INVENTION: CITRA-INTERACTING PROTEINS AND METHODS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

RESULT 12
US-10-045-202-4
: Sequence 4, Application US/10045202
: Publication No. US20030040461A1
: GENERAL INFORMATION:
: APPLICANT: Bristol-Myers Squibb
: TITLE OF INVENTION: MODULATORS OF BRUTON'S TYROSINE KINASE AND BRUTON'S TYROSINE KINASE INHIBITORS FOR THE TREATMENT OF B-CELL LYMPHOMAS AND USE IN THE TREATMENT OF B-CELL LYMPHOMAS
: TITLE OF INVENTION: PREVENTION OF OSTEOPOROSIS AND RELATED DISEASE STATES
: FILE REFERENCE: D0032 NP
: CURRENT APPLICATION NUMBER: US/10/045, 202
: CURRENT FILING DATE: 2002-06-27
: PRIOR APPLICATION NUMBER: 60/242,471
: PRIOR FILING DATE: 2000-10-23
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 659
: TYPE: PRT
: ORGANISM: Murine
: US-10-045-202-4

```

Query Match	6.9%	Score 140;	DB 9;	Length 659;
Best Local Similarity	22.7%	Pred. No. 0.015;		
Matches	58;	Conservative	29;	Mismatches 71; Indels 98; Gaps 11;

QY	125	RSQNIKGDAVSRRKKNKIPLP	PRPRPLTLTKKQPLP	PRPEESSRRPLSQNHHTPR	EVQGM	184
		!:::!!	!:::!!	!:::!!	!:::!!	
Db	171	RNGSLKPPSSIRKTKKPLP	PRPEEDQILTK--PLPPEAA	-----	AP	210
QY	185	ISLRLDSEVLBA-EKVP	HNQ-----RKPESTHLE	NONTOELPIALISSFT	SNHS	238
		!:::!!	!:::!!	!:::!!	!:::!!	
Db	211	VSTSELKRVALLDYDMP	NANDLQIRKDE	YFILEESH--LPW	-----	WRA 254
QY	239	RDHGGMOPCSPORCQ	PASCPHENILPYITTS	KRPFPKRSDKDYOH	NEMWTIGE	298
		!:::!!	!:::!!	!:::!!	!:::!!	
Db	255	RDKKG-----	QEGYIPSNVYT-----	EAEDSTEM	EWMSKH	288
QY	299	QAVEEAFKRNKDKGS	FLVRDCSTKSKEE	PYLAAYENKYNVAKIR	FLERNQ	358
		!:::!!	!:::!!	!:::!!	!:::!!	
Db	289	SQAQGLLKQDEKKE	EGEFLIRDS	SKAGK-----	YTVSV--	325
QY	359	RGDEKFDSEVEDIE	HY 374			
		!:::!!				
Db	326	-----	DPQGVIRHY 334			

RESULT 14

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:33:12 ; Search time 14.3724 Seconds
(without alignments)
2515.001 Million cell updates/sec

Title: US-09-856-061-4

Perfect score: 2020

Sequence: 1 FQNFSLPKRNSWPRINSATG.....GLRGDEKFDVEDIEHYKN 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332.5	16.5	533	2	A56110
2	315.5	15.6	533	2	B56110
3	149	7.4	724	2	A38747
4	146.5	7.3	724	2	A38748
5	146	7.2	724	2	A38749
6	143	7.1	659	2	I49553
7	140	6.9	659	2	I49444
8	140	6.9	659	2	A45184
9	138.5	6.9	981	1	FQNVGM
10	136.5	6.8	841	2	A42254
11	131.5	6.5	485	2	T37550
12	131	6.5	723	2	B38749
13	127	6.3	663	1	TYWVRR
14	125.5	6.2	728	2	H59435
15	125	6.2	377	2	S08636
16	125	6.2	585	2	A46209
17	125	6.2	593	1	UN0805
18	125	6.2	595	1	A55651
19	125	6.2	597	1	A53593
20	123.5	6.1	1196	2	T23832
21	123	6.1	461	2	A57463
22	122	6.0	822	1	TYVHFE
23	121.5	6.0	1893	1	A40262
24	121	6.0	1171	2	T12956
25	120.5	6.0	893	2	G88551
26	120	5.9	593	2	JC5167
27	120	5.9	939	2	S28394
28	120	5.9	996	2	T47518
29	119.5	5.9	2649	2	T51023

30	119	5.9	480	2	JC7552
31	119	5.9	620	1	S33253
32	118.5	5.9	1392	2	T51947
33	118	5.8	546	2	S52314
34	117.5	5.8	289	2	T52354
35	117.5	5.8	667	2	T27672
36	116.5	5.8	1162	2	T49191
37	116	5.7	545	2	S52313
38	115	5.7	675	2	S60612
39	115	5.7	796	2	E96654
40	115	5.7	1290	2	A36466
41	114.5	5.7	625	1	A43030
42	114.5	5.7	630	1	T01380
43	114.5	5.7	741	2	T40095
44	114	5.6	517	2	A43807
45	114	5.6	1128	2	A43960

ALIGNMENTS

RESULT 1
A56110
tyrosine phosphoprotein SLP-76 - human
C:Species: Homo sapiens (man)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C:Accession: A56110
R:Jackman, J.K.; Motto, D.G.; Sun, Q.; Tanemoto, M.; Turk, C.W.; Peltz, G.A.; Koretz, J. Biol. Chem. 270, 7029-7032, 1995
A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with A:Reference number: A56110; MUID:95221245; PMID:7706237
A:Accession: A56110
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <OAC>
A:Cross-References: GB:U20158; MID:9806765; PIDN:AAC50135.1; PID:9806766
C:Genetics:
A:Gene: GDB:ICP2; SLP-76
A:Cross-References: GDB:1230199; OMIM:601603
A:Map position: 5q33.1-5qter
C:Superfamily: SH2 homology
C:Keywords: phosphoprotein
F:422-520/Domain: SH2 homology <SH2>

Query Match 16.5% Score 332.5; DB 2; Length 533;
Best Local Similarity 25.3%; Pred. No. 1e-15;
Matches 113; Conservative 49; Mismatches 128; Indels 157; Gaps 16;

QY 20 GQYQKPKPLDWRNFAVLGAKGSHDDDDYDDELRLMEETWQSIKILPAPRIKSEYA 79
DB 126 GYVESN-----EEEAAYED-----DADYEPSPSNDDEALQN-SILPAKFPNS--- 169
QY 80 DTHYEVAMDPLDTRTSISIGQPTWNTQRLERVDKPISRDVSNQIKGDASVKNK 139
DB 170 NSMY-----IDRPPSGKTPQGP----- 186
QY 140 IPLPPRLITLP-----KKYQPLP-----EPSSR----- 166
DB 187 -PVPQPMALPLPPPPAGNNHSDPLPPQTNHPEPSRHHKAKLAPASIDNSTKPELDR 245
QY 167 -----PPLSQRTFP-----EVOGMPQSILDLSEVLEAEKV 199
DB 246 SLAPFREPFTLGKKRPFSDKPSIPAGRSGLHLPKIQKPLPPTTERHRRSSPLPGKKP 305
QY 200 P-----HNRKPESTHLEKNQT-----QETPLAISSSST 230
DB 306 PVPKHWGPDRENDSDVHQRPLPQALLPMSNTFSPSRSTKPSPMNPLPSHMGAP 365
QY 231 TSNHSVQNDHRCGMQPCSPQ--RCOPPASCSPEHETILPYKTTSMWPPPKRSDDKRDVQ 288
DB 366 ENSSSF---PQASSLPPTYSQGPSNRPPIRAGKRNFPPLP--NKRPPSP--AEENSLN 419
QY 289 NMYTIGYSRQAVAEAFMKENKDGSEFLVRDCSTKSKKEEYVLAVFENKVMYKIRFLER 348

Db 420 EEWVSYITPREAALARKINODGFILVRDSSKTTTPYLVLYLKQKVNIGIRYCKE 479
 QY 349 NQFALGTGLGDEKEDSVEDIIIEHYK 375
 Db 480 SQVYLGTGLGKEDFLSVSDIIDYFR 506

RESULT 2

B56110
 tyrosine phosphoprotein SLP-76 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999

C:Accession: B56110

R:Jackson, J.K.; Motto, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretzky, J. Biol. Chem. 270, 7029-7032, 1995

A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with C

A:Reference number: A56110; MUID:95221345; PMID:7706237

A:Accession: B56110

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-533 <IAC>

A:Cross-references: GB:020159; NID:9806767; PIDN:AAC52189.1; PID:9806768

C:Superfamily: SH2 homology

C:Keywords: phosphoprotein

F:422-520/Domain: SH2 homology <SH2>

Query Match 15.6%; Score 315.5; DB 2; Length 533;
 Best Local Similarity 30.0%; Pred. No. 1.5e-14;
 Matches 121; Conservative 44; Mismatches 169; Indels 69; Gaps 18;

QY 20 GOYRMNPKPLDMERNFAVLDGAKGSHDDYDDPELMEETWOSIKILPARPI--KESE 77
 Db 126 GDYESPN-----EEQALVDDAA-----DYEPSPNNEALOS-SILPNSFHNTSM 172
 QY 78 YAD-THYFKVAMDPL-PLDTRTISISIGPTWNTQTLERVDKPISRDVSQNIKGDASY 135
 Db 173 YIDRPTGKVSQPPVPLRPKPALP-PLPTGRNHSPLSPHPNHEPSRSGNKK----- 226
 QY 136 RKNKILPL-----PPRLTLTKKTYPL-----PPPESSRPLSRHNFPEVQ-GM 181
 Db 227 -TALPLASIDRSRTKPLDRLSLAPDRREPFLTGKPKPSDPAFLGRH-LPKIQKPL 284
 QY 182 PSQSLRLDSEVL-----EAEKVPNORKPESTHLE---NONTQELPLAISSSEFTSN 233
 Db 285 PPADNRHNRNRLGPRVYTRKRSVRHGRGPRRENDEDVQORPLPQSLSLMSSTNTPS 344
 QY 234 HSYQ-----NRDRGGMQPCS--PQRCQPPASCP--HENTLPEKYTS 272
 Db 345 RSVQSSKNTPPLAHMPCAFESNIGFQOSASLPPYFGCGNRPPLRSEGRNLPPLVPN 404
 QY 273 WRPPPRKSKDKVOHNMWYIGESRQAVEAFKKNKDSFLVRDCTSKKEEYVLAV 332
 Db 405 -RPPPSGEEETPLDEDMVSYITRPREAALARKINDGFLVVDSSKKTANNPYVLAV 463
 QY 333 FYENKVVNWKIRFLERNQOFGALGTGLGDEKEDSVEDIIIEHYK 375
 Db 464 LYKDKVNIQIRQESQVYLTGTGLGKEDFLSVSDIIDYFR 506

RESULT 3

A38747
 phosphatidylinositol 3-kinase (EC 2.7.1.-) 85k chain - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 16-Jul-1999

C:Accession: A38747

R:Escobedo, J.A.; Navarokasattusas, S.; Kavanaugh, W.M.; Milfay, D.; Fried, V.A.; William Cell 65, 75-82, 1991

A:Title: cDNA cloning of a novel 85 kD protein that has SH2 domains and regulates binding

A:Reference number: A38747; MUID:91191564; PMID:1849460

A:Molecule type: mRNA

A:Residues: 1-724 <ESC>

A:Cross-references: GB:M60651
 C:Comment: This protein binds a phosphotyrosine-containing sequence of ligand-activated phosphatidylinositol at position 3 of the inositol ring.
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:333-428/Domain: SH2 homology <SH2A>
 F:624-718/Domain: SH2 homology <SH2>

Query Match 7.4%; Score 149; DB 2; Length 724;
 Best Local Similarity 21.7%; Pred. No. 0.01;
 Matches 76; Conservative 39; Mismatches 96; Indels 140; Gaps 17;

QY 141 PLPPRPLITPKKYQPLPEPPSSR-----PLSQHRTPEVQGMPSQSLR 188
 Db 85 PTPKPRP-----PR---PLPVPAGSSKTPADTEQALPLPDLAEQFAPRPD-APLLIKLL 136
 QY 189 DLSEY--LEAKVPNORKPESTHLE-----NONTQELPLAISSSEFTSNHVNOR 239
 Db 137 EAIEKKGLECSLYRTQSSNPDELRLDCCDAASVDLEMIDVHLADAFKRYLADLPN- 195
 QY 240 DHRGGMQPCSP-----ORCQPPASC-----SPHENILPKYK----- 270
 Db 196 -----PVTFVAVYNEMSLAQELQSPEDCTQLKLLIRLPNIPHOQWLTLQYLKHF 248
 QY 271 -----TSW----- 273
 Db 249 KLSQASSKNLNAVRLSIFSPVLEFRPFAASDNTHEILKAILLISFENNERQPARALP 308
 QY 274 -RPPPP-----KSDRKDVONEMWYIGESRQAVEAFKKNKDSFLVRDCTSKKEE 326
 Db 309 PKPKPTTVANNMNNMSLQDAEYMWGDISREVENK-LRDTADGFLVVDASDKMHGD 367
 QY 327 PYVLAVFENKVVNWK-IRFLERNQOFGALGTGLGDEKEDSVEDIIIEHYK 376
 Db 368 -YTLTL--RRGNNKLIKIFHRDQKYGFSDDL-----TENSVELINHYRN 410

RESULT 4

A38748
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85k chain - human

C:Species: Homo sapiens (man)

C>Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 16-Jul-1999

C:Accession: A38748; S28402

R:Skolnik, E.Y.; Margolis, B.; Mohammadi, M.; Lowenstein, E.; Fischer, R.; Drepps, A. Cell 65, 83-90, 1991

A:Title: Cloning of P13 kinase-associated p85 utilizing a novel method for expression

A:Reference number: A38748; MUID:91191565; PMID:1849461

A:Accession: A38748

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-724 <SKO>

A:Cross-references: GB:M61906

R:Panayotou, G.; Bax, B.; Gout, I.; Federwisch, M.; Wroblewski, B.; Dhand, R.; Fry, M. EMBO J. 11, 4261-4272, 1992

A:Title: Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2 domain

A:Reference number: S28402; MUID:93049176; PMID:1330535

A:Accession: S28402

A:Status: preliminary

A:Molecule type: protein

C:Genetics:

A:Gene: GDB:PIK3R1

A:Cross-references: GDB:127604; OMIM:171833

A:Map position: Sq12-Sq13

C:Superfamily: SH2 homology

C:Keywords: phosphotransferase

F:333-428/Domain: SH2 homology <SH2A>
 F:624-718/Domain: SH2 homology <SH2>

Query Match 7.3%; Score 146.5; DB 2; Length 724;
 Best Local Similarity 24.7%; Pred. No. 0.015;
 Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;


```

QY 176 PEVOGMPQISLRDISELEAEKVPHNQ-----RKRESHILLNENOTCET 220
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 PEVOSSSEYIOL--LKKILIRSPSIDPHOYWTLQYLILKHFELLSQISSKNLNAKLSTIF 268

QY 221 -PLAISSSSFTTNSH-----SVONDRHRCGMOPCSPORCOPASCSPHENILPY 268
    : : : : : : : : : : : : : : : : : : : : : : : :
Db 269 SPMLFRFSAASDNTENLIKVIETLISFEMNERQ-----PAPALP----- 309

QY 269 KYTSMRPFPR-----RSDRKDVQHNEMWTIGEYSKQAVEAFMKENKQSEFLVDCSTK 322
    : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 -----KPKRPPTVANGNNNNMNSLONAEWGTDISREVEYNKR--LBDTADGTFVLVADASTK 363

QY 323 SKEPPYLAAYEYENKVYVWK--IRFLERNQOFALGTGLGDEKFPDSEVDIIEHYN 376
    : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 MHGD-YILLTL--RKGGNNKLKIFHRDGKTFGSDPL-----TFSSVELINHYRN 410

```

RESULT 5
A38749
3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain alpha - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 05-Nov-1999
C:Accession: A38749
R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, A
Cell 65, 91-104, 1991
A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine k
A:Reference number: A38749; MUID:91191567; PMID:1707345

A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-724 <OTS>
A: Cross-references: GB:M61746; GB:M61745; NID:9163476; PDB:AAA9511.1; PID:9163477
C: Superfamily: SH2 homology
C: Keywords: phosphotransferase
E: 333-428/Domain: SH2 homology <SH2A>
E: 624-718/Domain: SH2 homology <SH2>

Query Match	7.2%	Score 146	DR 2	Length 724
Best Local Similarity	22.4%	Pred. No. 0.016		
Matches 77; Conservative	49;	Mismatches 103;	Indels 114;	Gaps 16

```

Oy 136 RKNKPLEPPPLLTLPKYYOLRPPPESS--PPLSOHHPFV-----178
Db 80 KKSIPTRKPPRPPLVPAPRPSTLEADSEQASTLEDLAQAPRPVPAPPLIKLVEAI 139
Oy 179 --QGMPQISLR-----DLSEVLAEKVP-----HNQRPESTHLENOQT 220
Db 140 EKKGECSFTLYRQSSNPALRQLDODTASLDLEMDVHVLADAFKRRLDLPN--PVI 198
Oy 221 PLAISSSFT-----TSNSHV 236
Db 199 PVAVSELSILAPVEQSEERYIOLKLIRSPTRHQYWLTLQYLKHNFKRLQSTSSKNL 258
Oy 237 QNRDHGGMQPCSPQRCOPAPSCSPH--ENLP-----YKYSWR-----PPPKR-- 260
Db 259 LNA--RVLSLFSPLFFFPFAASENEHLKIIELISTEMNEROPALPRKPRKPT 316
Oy 281 -----SDRKVOHNEHIGEYSQAVAEAMKENKGSLFVLRDQSTKKEEPRVLAVFY 334
Db 317 VANNGMNNMNSLDQAEWWDGISNEBVENVK--LRDTAGCTFLVYRNASRKMNGD--YTLL-- 372
Oy 335 EKKYVNVK--IRPLERNOQFALGTGLRGDEKDSVEDIIEHKN 376
Db 373 --RKGGNNKLIRKIFHRDGYGFSDDL---TNSVVELIHNHRN 410

```

RESULT 6
149553
protein-tyrosine kinase (EC 2.7.1.112) emb - mouse
N:Alternate names: B-cell progenitor cytoplasmic tyrosine kinase; Bruton agammaglobulinemia
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Feb-2000

C:Accession: I49553; B45184; JN0471; PC4366
R:Sidecars, P.: Muller, S.; Shiels, H.; Jin, H.; Khan, W.N.; Nilsson, L.; Parkinson, I
Ulos, K.G.
J:Immunol. 153, 5607-5617, 1994
A:Title: Genomic organization of mouse and human Bruton's agammaglobulinemia tyrosine
A:Reference number: I49553; MUID:95081608; PMID:7989760
A:Accession: I49553
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-659 <SID>
A:Cross-references: GB:I29788; NID:9625143; PIDN:AAA66943.1; PID:9625144
R:Tsukada, S.; Saffran, D.C.; Rawlings, D.J.; Parolini, O.; Allen, R.C.; Klisak, I.;
Cell 72, 279-290, 1993
A:Title: Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked
A:Reference number: A45184; MUID:93145329; PMID:8425221

A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-122,'P',124-659 <TSU>
A:Cross-References: GB:108967; NID:g192233; PIDN:AAA37316.1; PID:g192234
A:Experimental source: 70z/3 pre-B cell line
A:Note: sequence extracted from NCBI backbone (NCBI:123834) and corrected to corres
R:Yanada, N.; Kawakami, Y.; Kimura, H.; Fukunachi, H.; Balser, G.; Altman, A.; Kato, '
Biochem. Biophys. Res. Commun. 192, 231-240, 1993
A:Title: Structure and expression of novel protein-tyrosine kinases, Emb and Emt, in
A:Reference number: JN0471, MUID:93236578; PMID:8476425
A:Accession: JN0471
A:Status: nucleic acid sequence not shown

A:Residues: 1-66, 'A', 68-197, 'W', 198-449, 451-659 <YAM>
A:Cross-references: GB:LI0627; NID:913016
A:Note: this translation is not annotated in GenBank entry MUSEMBX, release 116.0
A:Kojima, T., Fukuda, M., Watanabe, Y., Hamazato, F., Mitsuhashi, K.
Biochem. Biophys. Res. Commun. 236, 333-339, 1997
A:Title: Characterization of the pleckstrin homology domain of Btk as an inositol po
A:Reference number: PC4366; MUID:97382431; PMID:9240435
A:Accession: PC4366
A:Molecule type: mRNA
A:Residues: 1-165 <KOJ>
C:Comment: this protein specifically recognizes the inositol 1,3,4,5-pentakisphospha
e molecule to the membrane.

A:Gene: emb; Btk
C:Superfamily: protein-tyrosine kinase tec; pleckstrin repeat homology; protein kinase
D:Keywords: ATP; phosphoprotein; phosphotransferase; tyrosine-specific protein kinases
E:1-131/Domain: pleckstrin repeat homology
F:221-269/Domain: SH3 homology <SH3>
F:281-377/Domain: SH2 homology <SH2>
F:400-658/Domain: protein kinase homology <KIN>
F:7408-7416/Region: protein kinase ATP-binding motif
F:7551/Binding site: phosphate (Tyr) (covalent) status predicted

[illegible]

Db 326 -----EPGVIRHY 334

RESULT 7

149444

SH3 binding protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999

C:Accession: I49444

R:Ren, R.; Mayer, B.J.; Cicchetti, P.; Baltimore, D.

Science 259, 1157-1161, 1993

A:Title: Identification of a ten-amino acid proline-rich SH3 binding site.

A:Reference number: I49444; MUID:93174278; PMID:8438166

A:Accession: I49444

A:Molecule type: mRNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-559 <RES>

A:Cross-references: GB:I49443; NID:9293267; PIDN:AAA37121.1; PID:9293268

C:Superfamily: pleckstrin repeat homology

F:25-128/Domain: pleckstrin repeat homology <PLK>

F:201-210/Region: proline-rich SH3 binding

Query Match

Best Local Similarity 6.9%; Score 140; DB 2; Length 559;

Best Local Similarity 19.7%; Pred. No. 0.031;

Matches 93; Conservative 54; Mismatches 156; Indels 168; Gaps 18;

56 LRMEWTQSIKILPAPRKESFYADTHYFKVAM-----TPRLPD 95

Db 82 MRAEETSNVFPFKITHSKHRTWFSASSEDEKRSMAFVRREIGHHEKKEPLD 141

Qy 96 TRTSISIQPTWNTQRLERVDKPSRDVRSQNGKASVKKKIP--LPFRP----- 147

Db 142 TSDSS-----DTDSFGAVERPIDISLSYPMNEDYEHEDDDSYLEPSPMKLE 195

Qy 148 -LITPKKYQPLPEPSSRPPLS---QRHTF-----PEVOGMSQIL-RDL 190

Db 196 DALTYPPAYPP-PPVPPVRKPAFSDLPRAHSTSKSPSPPLPPPPKGLPDTGAPEDA 254

Qy 191 SEVLEAEKVPINQRKESTHLENQNTQEIPLAIS----- 225

Db 255 KDALGLRVEFGQLRVPAPRRNSDPMSNVPLVRLKHPCRDSVNGLEPMTGHGTS 314

Qy 226 --SSFTTSHNSVQNRD-----HRGKQP-----CSPQRCQ-- 254

Db 315 SVSSSTTAAVATSRNCDLKSFLHSSRGPPSEPPVPANKPKFLKIAEPPSPRAKFA 374

Qy 255 --PRASCSPHENILPYKXTSMRPPPRKSDRKDVH----- 288

Db 375 PVPVPAAPRPVQKMPKMPKPAVPAVLPREPNTPLHLQSPPDGSGFSGFEKARQPSQ 434

Qy 289 -----NEWYIGYSRQAVEAEAFMK-----ENKGSFLVDCSTKSKSKE 325

Db 425 ADTGEDEDEDEKVPPLNSVFNVTESCEVERLFKATDPGCEPDGLXCIRNSTKSGK 494

Qy 326 EPPYLAAYE--NKYVANKIRFLERNOQFALGTGLRDEKDSVEDITIEHY 374

Db 495 --VLVWDESSNKNVRYRI--FEKDSKFY----LEGEVLPAVSQWVEHY 536

RESULT 8

A45184

protein-tyrosine kinase (EC 2.7.1.112), nonreceptor type, BTK - human

N:Alternate names: alk; B cell progenitor kinase; Bruton agammaglobulinemia tyrosine kit

C:Species: Homo sapiens (man)

C:Date: 30-Apr-1993 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000

C:Accession: I37212; I54541; I39457; S28912; A45184

R:Ohle, Y.; Haire, R.N.; Litman, R.T.; Fu, S.M.; Nelson, R.P.; Kratz, J.; Kornfeld, S.J.

Proc. Natl. Acad. Sci. U.S.A. 91, 9066-9066, 1994

A:Title: Genomic organization and structure of Bruton agammaglobulinemia tyrosine kinase

A:Reference number: I37212; MUID:94377492; PMID:8090769

A:Accession: I37212

A:Status: nucleic acid sequence not shown; translation not shown; not compared with c

A:Molecule type: DNA

A:Residues: 1-659 <OHT>

A:Cross-references: EMBL:U10087; NID:q126232; PIDN:AAB60639.1; PID:9517438

A>Note: Only Intron-exon junctions are shown

R:Rohter, J.; Parolino, O.; Belmont, J.W.; Conley, M.E.

Immunogenetics 40, 319-324, 1994

A:Title: The genomic structure of human BTK, the defective gene in X-linked agammaglo

A:Reference number: I54541; MUID:95012452; PMID:7927535

A:Accession: I54541

A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-659 <ROH>

A:Cross-references: GB:L1572; NID:9632958; PIDN:AAA61479.1; PID:9632960

R:Hagemann, T.L.; Chen, Y.; Rosen, F.S.; Kwan, S.P.

Hum. Mol. Genet. 3, 1743-1749, 1994

A:Title: Genomic organization of the Btk gene and exon scanning for mutations in pat

A:Reference number: I39457; MUID:95152493; PMID:7880320

A:Accession: I39457

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-278, 281-659 <HAG>

A:Cross-references: EMBL:U13433; NID:9575688; PIDN:AAC51347.1; PID:9575690

R:Vettrle, D.; Vorechovsky, I.; Sideras, P.; Holland, J.; Davies, A.; Fliinter, F.; Ham

Nature 361, 226-233, 1993

A:Title: The gene involved in X-linked agammaglobulinemia is a member of the src fam

A:Reference number: S28912; MUID:93140868; PMID:8380905

A:Accession: S28912

A:Molecule type: mRNA

A:Residues: 1-659 <VER>

A:Cross-references: GB:X58957; NID:9312466; PIDN:CAA41728.1; PID:9312467; GB:U78027;

R:Tsukada, S.; Saffran, D.C.; Rawlings, D.J.; Parolinski, O.; Allen, R.C.; Klasek, I.;

Cell 72, 279-290, 1993

A:Title: Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-link

A:Reference number: A45184; MUID:93145329; PMID:8425221

A:Accession: A45184

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-19, 'R', 21-190, 'A', 199-206, 'T', 208-210, 'I', 212-218, 'E', 220-238, 'E', 240-2

A:Cross-references: GB:U78027; GB:I35265; NID:92281904

A:Experimental source: erythroleukemia cell line K562

A>Note: sequence extracted from NCBI backbone (NCBI:123835)

C:Genetics:

A:Gene: GDB:BTK; AGMX1; IMD1

A:Cross-references: GDB:120542; OMIM:300300

A:Map position: Xq21.33-Xq22

A:Introns: 47/3; 80/3; 103/3; 131/1; 174/1; 196/3; 259/2; 280/2; 298/3; 325/2; 368/1;

C:Superfamily: protein-tyrosine kinase; tyrosine-specific protein kinase

C:Keywords: ATP; phosphotransferase; pleckstrin repeat homology; protein kinase

F:2-131/Domain: pleckstrin repeat homology <PLK>

F:221-269/Domain: SH3 homology <SH3>

F:281-377/Domain: SH2 homology <SH2>

F:408-658/Domain: protein kinase homology <KIN>

F:408-416/Region: protein kinase ATP-binding motif

Query Match

Best Local Similarity 6.9%; Score 140; DB 2; Length 659;

Best Local Similarity 22.7%; Pred. No. 0.038;

Matches 58; Conservative 29; Mismatches 71; Indels 98; Gaps 11;

Qy 125 RSQNTKGASVRRKNIPLPPRPRLITLPKKTQPLPEPSSRPPLSQHRTPEVOGMPQ 184

Db 171 RNSGLKPCSSHRKTKKPLPPPEEDQILK--PLPEPAA-----Ap 210

Qy 185 ISLRDLSEVLA-EKVPNQ-----RKPESTHLENQNTQEIPLAISSSFTTSHNSVON 238

Db 211 VSTSEIKVAVALLDYPNANMDLQKDEFFILESN--LPW-----WRA 254

Qy 239 RDHKGMPQSPORQCPASCPHENILPYKYTSMRPPPRKSDRKDVQHNEWYIGEYSR 298

Db 255 RDKNG-----QEGYIPSNYVT-----EAEDSIEMEWYSKIMTR 288

Qy 299 QAVEAEAPKKNKDGSLVRDCSTSKSEPPYLAAYENKVVNVIKIRFLERNOQFALGTGL 358

Db 289 SOAEOLKQEGEGFIVRDSSKAGK-----YTVSV-----FAKSTG- 325
 QY 359 RGDEKFDVEDIIEHY 374
 Db 326 -----DPQGVIRHY 334

RESULT 9

FOMVGM

gag-abl polyprotein - Abelson murine leukemia virus

N:Contains: amino end of core shell protein p30; core protein p15; inner coat protein p1

C:Species: Abelson murine leukemia virus

A:Note: host Mus sp. (mouse)

C:Date: 14-Nov-1983 #sequence_revision 09-Sep-1994 #text_change 11-Jun-1999

C:Accession: A03931; A00627; A93955

R:Reddy, E.P.; Smith, M.J.; Strinivasan, A.

Proc. Natl. Acad. Sci. U.S.A. 80, 3623-3627, 1983

A:Title: Nucleotide sequence of Abelson murine leukemia virus genome: structural similar

A:Reference number: A93955; MUID:83221648; PMID:6304726

A:Accession: A03931

A:Molecule type: DNA

A:Residues: 1-981 <RED>

A:Cross-references: GB:J02009; NID:9331887; PIDN:AAA46471.1; PID:9331888

A:Note: the authors translated the codon GGA for residue 186 as Glu

R:Reddy, E.P.; Smith, M.J.; Strinivasan, A.

Proc. Natl. Acad. Sci. U.S.A. 80, 7372, 1983

A:Reference number: A93980

A:Contents: annotation; erratum, residues 588-746

C:Genetics:

A:Gene: gag-abl

C:Superfamily: Abelson murine leukemia virus gag-abl polyprotein; protein kinase homolog

C:Keywords: ATP; core protein; oncogene; phosphotransferase; polyprotein; transforming F

F:1-131/Product: core protein p15 #status predicted <p15>

F:132-215/Product: inner coat protein p12 #status predicted <p12>

F:216-235/Region: amino end of core shell protein p30

F:248-338/Domain: SH2 homology <SH2>

F:361-621/Domain: protein kinase homology <kin>

F:369-377/Region: protein kinase ATP-binding motif

F:392/Active site: Lys #status predicted

Query Match

Best Local Similarity 6.9%; Score 138.5; DB 1; Length 981;

Matches 62; Conservative 40; Mismatches 90; Indels 79; Gaps 16;

QY 143 PPP--RPLTLPPKYQPLP-----EPSSRPPLSQRH--TFPEVQMPQSISLRD-- 189

Db 93 PPPWVPEVH-PKPPPLPPSAPSLPEBPPLSTPPRSSLYPLTPELGAKPKPQVLSDSG 151

QY 190 --LSEVLEAKYPHNRKRKESTHLENONTQELPLAISSSFTTSMHSVQNRDHRGC-MQ 246

Db 152 GPLIDLTEDPPPRPRPPS-----DRGNGGEAT 183

QY 247 PC-----SPQ-----RCQPPASCSPHENILPYK-----YTWMPPPPKRSR----- 283

Db 184 PGEAPDEPSPMSRLRGREPPVADSTQAFPLRTGGNGQLQY--PF--SSSDLYITTV 240

QY 284 KVVQHNWETIGEYSROAVEAFMKENKDGSEFLVRDCTSKSEEPYLAFTENKYVNVKI 343

Db 241 NSLEKSHWYHGFSVMAA-EYLLSSGINSFLVRE--SSSPGQGISLRYEGRVYHYRI 297

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQAFALGTGLGDEKFDSEVDIIEHY 374

A:Title: corkscrew encodes a putative protein tyrosine phosphatase that functions to

A:Reference number: A43254; MUID:92346711; PMID:1638629

A:Accession: A43254

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-841 <PER>

A:Cross-references: GB:M94730; NID:9157144; PID:9157145

A:Experimental source: embryo

A:Note: sequence extracted from NCBI backbone (NCBIN:109964, NCBIPI:109965)

C:Genetics:

A:Gene: FlyBase:csw

C:Superfamily: FlyBase:Pg0000382

C:Keywords: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-pho

F:6-101/Domain: SH2 homology <SH2>

F:111-203/Domain: SH2 homology <SH2>

F:252-634/Domain: protein-tyrosine-phosphatase homology <PRP>

F:583/Active site: Cys (phosphocysteine intermediate) #status predicted

F:589/Binding site: substrate phosphate (Arg) #status predicted

Query Match

Best Local Similarity 6.8%; Score 136.5; DB 2; Length 841;

Matches 54; Conservative 39; Mismatches 67; Indels 71; Gaps 9;

QY 170 SQRTPEVQMPQSISLRDSEVLEAEKVPYHNC-----RK 206

Db 3 SRWPFHTISG-----IEAKLDEQGFDSFLARLSSNGCAFTLSVRRN 49

QY 207 ESTHLENONTQELPLAISSSFTTSMHSVQNRDHRGCKQSPQ--RCQPPASCSPHEN 264

Db 50 EYTH-IKIONNDPFDLYGGEKFAFLPELVQYMMENGELKRNQAIETLKOPLICA---- 104

QY 265 ILPYKTSRPPPKRSKRDVQHNWYIGESFOAVEAFMKENKDGSEFLVRDCTSK 324

Db 105 -----EPFTR-----WFGNLSGKAELILIERKNGSFLVRE--SQSK 142

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 143 PCDFVLAVRTDVKVTHVMIRW--QDKKYDVG---GGSEFGFLSELIDHYK 187

QY 325 EEPYLAVEYENKYNVNIREFLRNQAFALGTGLGDEKFDSEVDIIEHY 375

Db 260 NEIITKDTPLKDKYSKAPALLPQKPKVS---KQIYOQVSVFSTGKKIESQSLNLTDT 316
 Qy 218 Q-ELPPLAIISSSTFTSNHSONRHRGMPQPCQRCQPCAPSPHENILPYKTTSM--- 273
 Db 317 DIETPLKGS-----ELYSEDFKPNVDPVKIQOI-----LHKONKIIIEKMIOSIRI 364
 Qy 274 -----RPPFKRSRDKVDONHMEYIGESRQAVAEATMKENKDGSLVRCST 321
 Db 365 SKNEVQRQLDQERHAELETAKIENNRFLTKRRRKA-REAIQKIDN-----LKDLSV 418
 Qy 322 KSKEEPLYLAVYENKYNVAKIRLEHNOQFALGTGDEKFD 365
 Db 419 ---QELFIPSERELKYELK-----RKDEKLD 443

RESULT 12

B38749
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain B - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 16-Jul-1999
 C:Accession: B38749
 R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, A.
 Cell 65, 91-104, 1991
 A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine k
 A:Reference number: A38749; MUID:91191567; PMID:1707345
 A:Accession: B38749
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-723 <OTS>
 A:Cross-references: GB:M61745; GB:M61746
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:325-420/Domain: SH2 homology <SH2A>
 F:617-706/Domain: SH2 homology <SH2>

Query Match 6.5%; Score 131; DB 2; Length 723;
 Best Local Similarity 21.5%; Pred. No. 0.18;
 Matches 81; Conservative 50; Mismatches 118; Indels 128; Gaps 18;

Qy 101 STG-QPTWNTOTLERVDKPSRDVRSNKGASVRRKNTPLPPRPPLTLTPKKYOLP 159
 Db 53 SVGMWPGINERT-R-QRGDFPCT---YVEFLGPVALAR---PGPRPRGRPRPLPARPRDGP 104
 Qy 160 PEPPSRPPLSQRHTFPVQGMPSQISLRDSEVLEAKVPHNORKPRESTHLENQOE 219
 Db 105 PEGULTLPDLPEQSPDPV-APPLVVL-----VEATERTGLDSYRPPRAVRTDWSLSD 158
 Qy 220 IP-----LAISSSFTTS-----NHSYON 238
 Db 159 VEQMDAALSDGVKGLLALPARLVTPRAAEHRALEAGVPALEPTLPLHNAITL 218
 Qy 239 R---DHGGMQPCSPQ-----RCQPPASCP----- 261
 Db 219 RFLQHLGRVAGRAPAPAVRATGATGFPLLRAPPPSPPGAPDGTPTDFPALL 278
 Qy 262 -----HENILPYKYTSMRPPPKRS-----DRKDVONENYIGESYQA 300
 Db 279 VEKLQHELEQEVAPPL-----PPKPKPIPAATGLANGSPSLQDAEMVIGDISKEE 334
 Qy 301 VEAFAEMKENKDSFLVRDCTSKSEEPYVLAIFYENKYNV-KIRLEHNOQFALGTGLR 359
 Db 335 VNEK-LRPTPGDTFLVRNASSKIGE-YTLTL---RKGNKKLKVHFRDGHYGSSEPL- 388
 Qy 360 GDEKFDSEVEDIEHYKN 376
 Db 389 ---TFCSYVDLITHYRH 402

RESULT 13

TWVWR
 protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rashed

C:Species: feline sarcoma virus
 A:Note: host Felis sp. (cat)
 C>Date: 27-Nov-1985 #sequence_revision 26-May-1995 #text_change 31-Mar-2000
 C:Accession: A00653; A03937
 R:Naharro, G.; Robbins, K.C.; Reddy, E.P.
 Science 223, 63-66, 1984
 A:Title: Gene product of v-fgr onc: hybrid protein containing a portion of actin and
 A:Reference number: A00653; MUID:84097512; PMID:6318314
 A:Accession: A00653
 A:Molecule type: DNA
 A:Residues: 1-663 <NAH>
 A:Cross-references: GB:X00255; GB:K01487; NID:961542; PIDN:CA025063.1; PID:961543
 A:Note: the authors translated the codon GAT for residue 14 as Glu
 C:Comment: This protein is synthesized as a gag-fgr polypeptide.
 C:Genetics:
 A:Gene: fgr
 C:Superfamily: feline sarcoma virus protein-tyrosine kinase fgr: protein kinase homol
 C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; p
 F:1-118/Region: gag polypeptide similarity
 F:141-268/Region: actin similarity
 F:285-382/Domain: SH2 homology <SH2>
 F:402-660/Domain: protein kinase homology <KIN>
 F:410-418/Region: protein kinase ATP-binding motif
 F:432/Active site: Lys #status predicted

Query Match 6.3%; Score 127; DB 1; Length 663;
 Best Local Similarity 19.1%; Pred. No. 0.31;
 Matches 59; Conservative 44; Mismatches 88; Indels 118; Gaps 12;

Qy 142 LPPRPPLTLTKKYOLP-----PEPSSRPPLSQ----- 171
 Db 101 LPPKPTSLPQHPSPQPARALCPAVCRPRPLPLPTAAEEVVALVIDNGSGMKAG 160
 Qy 172 -----RHFPPEVQGM-----SQISLRDSEVLE----- 195
 Db 161 FAGDAPRAVFPSTVGRHGVAVGQKDSYGDQSRGLTLTKYPEHGIYTNMD 220
 Qy 196 -AEKVPNQRKPESTHLENO---NTOEIPLAISSTFTSNHSONRHRGMPQPCSPQ 251
 Db 221 DMEKIWH-----HMFYNELRVAPREHPLVLEA-----PLNPK 253
 Qy 252 RCQPPASCPHENILPYKYTSMRPPPKRSRDKVONHMEYIGESRQAVAEATMK- NK 310
 Db 254 -----ANREKMTQINFETFNIPSNVAVPDSIQAEVEWFGKIGRKDEROLLSGNA 305
 Qy 311 DGSFLVRDCTSKSEEPYVLA-----FYENKYNVIRFLEHNOQFALGTGDEKFD 365
 Db 306 RGAFLVRESST--TKGAVSLSRDMDARGDHVKHYKIRKIDG-----GYITTRQGFN 358
 Qy 366 SVEDIEHY 374
 Db 359 SVQELVQHY 367

RESULT 14

H59435
 phosphoinositide-3-kinase regulatory beta chain [imported] - huam
 C:Species: Homo sapiens (man)
 C>Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: H59435; A59436
 R:Volinia S.; Patrascu I. P.; Otsu M.; Hiles I.; Gout I.; Calzolari E.; Bernardi F.; Rooke L.
 Oncogene 7, 789-793, 1992
 A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinosit
 A:Reference number: H59435
 A:Accession: H59435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <VOL>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1
 R:Janssen, J.W.; Schleithoff, L.; Bartlam, C.R.; Schulz, A.S.
 Oncogene 16, 1767-1772, 1998
 A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta sub
 A:Reference number: A59436; MUID:98241181; PMID:9582025

A:Accession: A59436
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-728 <JAN>
A:Cross-references: GB:NP_005018; PID:g4826908; PIDN:NP_005018.1

Query Match 6.2%; Score 125.5; DB 2; Length 728;
Best Local Similarity 20.9%; Pred. No. 0.44;
Matches 72; Conservative 31; Mismatches 101; Indels 141; Gaps 14;

QY 141 PLP-----PPRLTLPKRYQPLPE--PESSRPL-----SQRHTPEV 178
DB 95 PLAPRPDGAPEPGLTP---DLPEQSPDVAAPLLVLAIERTGLDSESHYRPEL 150
QY 179 QGMPQSISLRLSEVLEAEKVPNNORKPESTHLENQTOEIPLAISSTFTSNHSVN 238
DB 151 PAPRTDMSLSDVDQMDTAALA-----DGKSFLLALPAPLVTPASAEAR 195
QY 239 RDRHGMQPCSPQ-----251
DB 196 RALREAGAPVPALEPPTLPALRALTLRFLQHLGRVARRAPALGPAVRALGATFGPLL 255
QY 252 RCGPPASCSP-----HENILPYKTSWRPPPKRS-- 281
DB 256 RAPPSPSPGAPDGPSESPDPFALLVKKLLQEHLEQEVAPPAL---PPKPPKAP 311
QY 282 -----DRKDVQNEWYIGESRQAVEAFMKENKDGFLVRDCSTKKEEPYLA 332
DB 312 APVYLANGSPSLQDAEWYWDGISREVENK-LRTPDGTFLVRDASSKIQGE-YTLTL 369
QY 333 FYENKYYNWK-IRLELRNOQFALGTGLRGDEKFDSEDIIEHYKN 376
DB 370 ---RKGNKMLIKVFRHDGHYGFSEPL---TFCSVVDLINHYRH 407

RESULT 15

S08636
NCK Protein - human
N:Alternate names: src-related protein
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 05-Nov-1999
C:Accession: S08636
R:Lehmann, J.M.; Riethmuller, G.; Johnson, J.P.
Nucleic Acids Res. 18, 1048, 1990
A:Title: Nck, a melanoma CDNA encoding a cytoplasmic protein consisting of the src homol
A:Reference number: S08636; MUID:90192089; PMID:2107526
A:Accession: S08636
A:Molecule type: mRNA
A:Residues: 1-377 <LEH>

A:Cross-references: EMBL:X17576; NID:g35014; PIDN:CAA35599.1; PID:g35015
C:Superfamily: SH3 homology; SH2 homology
F:9-56/Domain: SH3 homology <SH31>
F:113-160/Domain: SH3 homology <SH32>
F:197-247/Domain: SH3 homology <SH33>
F:282-371/Domain: SH2 homology <SH2>

Query Match 6.2%; Score 125; DB 2; Length 377;
Best Local Similarity 23.6%; Pred. No. 0.21;
Matches 33; Conservative 23; Mismatches 52; Indels 32; Gaps 4;

QY 236 VQNRDRHGMQPCSPQRCPPASCSPHENILPYKYTSWRPPPKRSRDRKDVQHNEWYIGE 295
DB 250 MNNPLTSLGLEPSPQ-----CDYIRPSLTGKFAG-----NPMYGGK 286
QY 296 YSRQAVEEAFMKENKDGFLVRDCSTKKEEPYLAFAVENKYYNWKIRLELRNOQFALG 355
DB 287 VTRHQAEMLNKGHGDFLIRDSSESPDVSFLKAGKNNKFKVOL-----KETVYCI- 341
QY 356 TGLRGDEKFDSEDIIEHYK 375
DB 342 ----GQRKFSMEIVEHYK 357

Search completed: April 21, 2003, 12:38:38
Job time: 18.3724 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:28:57 ; Search time 8.34525 seconds
(without alignments)
1868.739 Million cell updates/sec

Title: US-09-856-061-4

Perfect score: 2020

Sequence: 1 FQNFSLPKNRSMPRINSATG.....GLRGDEKFDVEDIEIHKYKN 376

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332.5	16.5	533	1	LCP2_HUMAN
2	315.5	15.6	533	1	LCP2_MOUSE
3	149	7.4	724	1	P85A_MOUSE
4	146.5	7.3	724	1	P85A_HUMAN
5	146	7.2	724	1	P85A_BOVIN
6	145	7.2	724	1	P85A_RAT
7	143	7.1	659	1	BRK_MOUSE
8	140	6.9	559	1	BRP2_MOUSE
9	140	6.9	659	1	BRK_HUMAN
10	136.5	6.8	845	1	CSW_DROME
11	136	6.7	561	1	3BP2_HUMAN
12	133.5	6.6	724	1	P85B_BOVIN
13	128	6.3	461	1	P55G_HUMAN
14	127	6.3	271	1	CRK_DROME
15	127	6.3	722	1	P85B_MOUSE
16	126.5	6.3	533	1	SOC6_MOUSE
17	126	6.2	461	1	P55G_BOVIN
18	125.5	6.2	728	1	P85B_HUMAN
19	125	6.2	377	1	NCK1_HUMAN
20	125	6.2	585	1	PTN2_MOUSE
21	125	6.2	593	1	PTN2_HUMAN
22	124.5	6.2	217	1	GRAP_HUMAN
23	124.5	6.2	217	1	GRAP_MOUSE
24	123.5	6.1	1196	1	ABL1_CAELI
25	123	6.1	461	1	P55G_MOUSE
26	122	6.0	722	1	P85B_RAT
27	122	6.0	822	1	PER_HUMAN
28	120.5	6.0	893	1	YH92_CAELI
29	120.5	6.0	1490	1	CRK7_HUMAN
30	120	5.9	847	1	VAV3_HUMAN
31	120	5.9	939	1	STR2_YEAST
32	119	5.9	620	1	ITK_HUMAN
33	119	5.9	1872	1	T2D1_HUMAN

ALIGNMENTS

34	116	5.7	517	1	PCR_MOUSE	P14234 mus musculus
35	115	5.7	675	1	BMX_HUMAN	P51813 homo sapien
36	115	5.7	1290	1	P1G1_HUMAN	P19174 homo sapien
37	114.5	5.7	625	1	ITK_MOUSE	O03526 mus musculus
38	114.5	5.7	630	1	TEC_MOUSE	P24604 mus musculus
39	114	5.6	1128	1	BEM3_YEAST	P32873 saccharomyc
40	114	5.6	1190	1	ZO2_HUMAN	O94d72 homo sapien
41	114	5.6	2440	1	NCR1_HUMAN	O75376 homo sapien
42	112.5	5.6	303	1	CRKL_HUMAN	P46109 homo sapien
43	112.5	5.6	878	1	VAV2_HUMAN	P52735 homo sapien
44	112	5.5	757	1	HT16_HYDAT	P53356 hydra atten
45	112	5.5	1290	1	P1G1_RAT	P10686 rattus norv

RESULT 1

ID	LCP2_HUMAN	STANDARD:	PRT:	533 AA.
AC	Q13094			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).			
GN	LCP2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=Leukemia;			
RX	MEDLINE=95221345; PubMed=7706237;			
RA	Jackman J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A.,			
RA	Koretzky G.A., Findell P.R.;			
RT	"Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells."			
RL	J. Biol. Chem. 270:7029-7032(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
CC	- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.			
CC	- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND Fyb.			
CC	- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES. HIGHLY EXPRESSED ALSO IN T CELL AND MONOCYTIC CELL LINES, EXPRESSED AT LOWER LEVEL IN B CELL LINES.			
CC	NOT DETECTED IN FIBROBLAST OR NEUROBLASTOMA CELL LINES.			
CC	- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION BY ZAP-70.			
CC	- SIMILARITY: CONTAINS 1 SH2 DOMAIN.			
CC	- DATABASE: NAME=PROT; NOTE=PROT 1:1-5(2000);			
CC	WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1118450040.g.htm"			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U20158; AAC50135.1; -			
DR	HSSP; BC016618; AAH16618.1; -			
DR	Genew; HGNC:6529; LCP2.			
DR	InterPro; IPR001660; SAM.			
DR	InterPro; IPR000980; SH2.			
DR	Pfam; PF00017; SH2; 1.			

DR PRODOM; PD000093; SH2; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS00001; SH2; 1.
 KW SH2 domain; Phosphorylation.
 FT DOMAIN 133 136 POLY-GLU.
 FT DOMAIN 198 201 POLY-PRO.
 FT DOMAIN 422 530 SH2.
 SQ SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;

Query Match 16.5%; Score 332.5; DR 1; Length 533;
 Best Local Similarity 25.3%; Pred. No. 1.3e-15;
 Matches 113; Conservative 49; Mismatches 128; Indels 157; Gaps 16;

DR 20 GQYRMMKPLDWERNAVALDGAHSDDDYDDPELMEETQSIKILPAPKESSEYA 79
 DB 126 GDYSPN-----EEELAPED-----DADYEPPSPNDEALON-SILPAKPPNS--- 169
 QY 80 DTHYFKVAMDTPLDTRTISIGQPTWNTQTRLEVDKPTISRDVSQNIKGASVRKNK 139
 DB 170 NSMY-----IDRPPSGKTPQP----- 186
 QY 140 IPLPPRLTLPL-----KKYQPLP-----EPSSR----- 166
 DB 187 -PVPQRPMAALPPRPAGRNHSPPLPPQTNHEERSRNRHKTAKLPAPSIDRSTKPLDR 245
 QY 167 -----PPLSORHTPP-----EVOGMPQISLRDSEVLEAEKV 199
 DB 246 SLAFDEPPTLGGKPPSPDKPSIPAGRSGLGHLPKIQKRPPLPTTHERHSSPLPKKP 305
 QY 200 P-----HNQKPESTHLENT-----OEIPLAISSSFT 230
 DB 306 PVPKGGCPDRENDVDHQRPLPQPLPLMSSTPPSRSTKPPMNPPLPSHMPGAFS 365
 QY 231 TSNISVONRDRHGMQCSQP--KQCPASCSPIENILPYKTYWRPFPKRSRKNVOH 288
 DB 366 ESNSSF---PQASLPPYFSGSPSNRPPIRAEGRNPLPLP-NKRPSP--AAEENSLN 419
 QY 289 NEWIGESROAVEAEFMKKNKDGSLVRODSTKSKKEPYLAVFYENKYNVIRFLER 348
 DB 420 EEWVSYITRPEALAKRINKQDSTFLVRDSSKTTINPIYLVNLYDKYNIQIRQKE 479
 QY 349 NQOFALGTGLGDEKFDSEVDIEHYK 375
 DB 480 SQVLLGTGLGKEDFLSVSDIIDYFR 506

RESULT 2

LCP2_MOUSE STANDARD; PRT; 533 AA.

ID LCP2_MOUSE
 AC 060787;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
 GN LCP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell lymphoma;
 RX MEDLINE=95221345; PubMed=7706237;
 RA Jackman J.K., Motta D.G., Sun Q., Tanemoto M., Turck C.W., Pelz G.A., Koretzky G.A., Finkel P.R.;
 RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells.";
 RL J. Biol. Chem. 270:7029-7032(1995).
 CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
 CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND Fyb.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES.
 CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; U20159; AAC52189.1; -
 DR HSSP; P12931; 1SHD.
 DR MGD; MG1:1321402; LCP2.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR PRODOM; PD000093; SH2; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS00001; SH2; 1.
 KW SH2 domain; Phosphorylation.
 FT DOMAIN 422 530 SH2.
 SQ SEQUENCE 533 AA; 60228 MW; 6DDB77782C2E60C3 CRC64;

Query Match 15.6%; Score 315.5; DR 1; Length 533;
 Best Local Similarity 30.0%; Pred. No. 1.9e-14;
 Matches 121; Conservative 44; Mismatches 169; Indels 69; Gaps 18;

DR 20 GQYRMMKPLDWERNAVALDGAHSDDDYDDPELMEETQSIKILPAPKESSEYA 79
 DB 126 GDYSPN-----EEELALVDAA-----DYEPSPNNEALQS-SILPNSFHTNMS 172
 QY 78 YAD-THYFKVAMDTPL-PLDTRTISIGQPTWNTQTRLEVDKPTISRDVSQNIKGASV 135
 DB 173 YIDRPPYKVSQOPVPPVPLRPKPLP-PLPTGRNHSPLSPPHNHEERSGSKNKK----- 226
 QY 136 RKNKPIPLP-----PRLPLTLPPKKYQPL-----PEEPSSRPPLSORHTPPEVQ--GM 181
 DB 227 -TAKLPAPSIDRSTKPLPLDSLAPLDREPTLGGKPPSPDKPSAPLAGRH-LRKIQKPL 284
 QY 182 PQSILSLDLSEVL-----EAEKYPHNQKPESTHLE--NONTQEIPLAISSSFTSN 233
 DB 285 PRAMDREHNERERLGPVYTRKPSVPRHGRPDRENDVDHQRPLPQPLPLMSSTPPS 344
 QY 234 HSYV-----NRDHRGMQCS--PQRCQPPASCSP--HENILPYKYS 272
 DB 345 RSYQSSKNTFPLAHMPGAFSESNIGFQGSASLPPYFSGPGNRPPLRSBGRNLPPLVPN 404
 QY 273 WRPFPPKRSRDKDVQHNWIGESROAVEAEFMKKNKDGSLVRODSTKSKKEPYLAV 332
 DB 405 -RPQPPSGEETPLDEWVSYITRPEALAKRINKQDSTFLVRDSSKTTANPIYLVN 463
 QY 333 FYENKYNVNIKIRELNQOFALGTGLGDEKFDSEVDIEHYK 375
 DB 464 LYKDKYNIQIRYQESQVLLGTGLGKEDFLSVSDIIDYFR 506

RESULT 3

P85A_MOUSE STANDARD; PRT; 724 AA.

ID P85A_MOUSE
 AC P26450;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase p85-alpha subunit) (Pcdins-3-kinase p85-alpha) (PI3K).
 GN PI3KRI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-BALB/C;
 RA MEDLINE=91191564; PubMed=1849460;
 RA Escobedo J.A., Navankasatutas S., Kavanaugh W.M., Miley D.,
 RA Fried V.A., Williams L.T.;
 RT "cDNA cloning of a novel 85 kd protein that has SH2 domains and
 RT regulates binding of PI3-kinase to the PDGF beta-receptor";
 RL Cell 65:5-82(1991).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M60651; AAA39886.1; -
 DR HSP: P23727; 1BFI.
 DR MGI: 97583; PI3Krl.
 DR InterPro: IPR001720; PI3K_kinase_P85.
 DR InterPro: IPR000198; RhogAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PI3KINASEP85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000056; SH3; 1.
 DR PRODOM: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 KW SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 3 79 SH3.
 FT DOMAIN 333 428 SH2.1.
 FT DOMAIN 624 718 SH2.2.
 SO SEQUENCE 724 AA; 83414 MW; DEAA6BE297CF07A CRC64;

Query Match 7.48; Score 149; DB 1; Length 724;
 Best Local Similarity 21.7%; Pred. No. 0.0055;
 Matches 76; Conservative 39; Mismatches 96; Indels 140; Gaps 17;

QY 141 PUPPPRLTLPKYQPLPEPSSR-----PPLSQRHPTPEVQGPSQSLR 188
 DB 85 PTPKPRP-----PR---PLPAPGSSKTEADTEQOALPLPLAEOFAPDPV-APPPLTKL 136
 QY 189 DLSEV---LEAEKYPHNOKEPSTHLE-----NONTQRIPLAISSSTTSNHSYQR 239
 DB 137 EALEKKGLESTLYRQSSNPALRLQDLDCDAASVDLEMDIVHVLADAKRIADLPN- 195
 QY 240 DHKGMQPCSP-----ORCQPPASC-----SPHENILPYRY----- 270
 DB 196 -----PVIPLAVYEMMSLAQELQSPEDICIDLLKRLPLNIPHCWTLTLYLKHFF 248
 QY 271 -----TSN----- 273

DB 249 KLSQASSKNLLNARVLSFISPPVLFPPAASSDNTFHLIAITLITSEWNERQAPALP 308
 QY 274 -RPPFP-----KRSDDRVOHNEWTGYEYSROAVEAEFKENKDGSEFLVRCSTSKSE 326
 DB 309 PKRPKPTTVANNSSNNMNNNSLQDAEMWYGDISRREVENK-LRDPADGFLVRDASTKMHGD 367
 QY 327 PYVLAVYEKKYVNVK-IRFLERQCALGTGLRGDEKFEVSVDIITHYKN 376
 DB 368 -YTLTL--RRKGNNKLKIKFIHRDGKYGFSDDL-----TFNSVELLNHYRN 410

RESULT 4
 ID P85A_HUMAN STANDARD; PRT; 724 AA.
 AC P27986;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
 DE p85-alpha subunit) (PtdIns-3-kinase p85-alpha) (PI3K).
 GN PI3KR1 OR GRB1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91191565; PubMed=1849461;
 RA Skolnik E.Y., Margolis B., Mohammadi M., Lowenstein E., Fischer R.,
 RA Drees A., Ullrich A., Schlessinger J.;
 RT "Cloning of PI3 kinase-associated p85 utilizing a novel method for
 RT expression/cloning of target proteins for receptor tyrosine
 RT kinases";
 RL Cell 65:83-90(1991).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 1-85.
 RX MEDLINE=96196433; PubMed=8648629;
 RA Liang J., Chen J.K., Schreiber S.L., Clardy J.;
 RT "Crystal structure of PI3K SH3 domain at 2.0 Å resolution";
 RL J. Mol. Biol. 257:632-643(1996).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 324-434.
 RX MEDLINE=96185448; PubMed=8599763;
 RA Nolte R.T., Eck M.J., Schlessinger J., Shoelson S.E., Harrison S.C.;
 RT "Crystal structure of the PI 3-kinase p85 amino-terminal SH2 domain
 RT and its phosphopeptide complexes";
 RL Nat. Struct. Biol. 3:364-373(1996).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 115-298.
 RX MEDLINE=97121392; PubMed=8962058;
 RA Musacchio A., Cantley L.C., Harrison S.C.;
 RT "Crystal structure of the breakpoint cluster region-homology domain
 RT from Phosphoinositide 3-kinase p85 alpha subunit";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:14373-14378(1996).
 RN [5]
 RP STRUCTURE BY NMR OF 1-79.
 RX MEDLINE=93208889; PubMed=7681364;
 RA Koyama S., Yu H., Dalgarino D.C., Shin T.B., Zydzinsky L.D.,
 RA Schneider S.L.;
 RT "Structure of the PI3K SH3 domain and analysis of the SH3 family";
 RL Cell 72:945-952(1993).
 RN [6]
 RP STRUCTURE BY NMR OF 91-104.
 RX MEDLINE=97121261; PubMed=8961927;
 RA Renzoni D.A., Pugh D.J., Siligardi G., Das P., Morton C.J., Rossi C.,
 RA Waterfield I.D., Campbell I.D., Ladbury J.E.;
 RT "Structural and thermodynamic characterization of the interaction of
 RT the SH3 domain from Fyn with the proline-rich binding site on the p85
 RT subunit of PI3-kinase";
 RL Biochemistry 35:15646-15653(1996).
 RN [7]
 RP STRUCTURE BY NMR OF 617-724.
 RX MEDLINE=96312955; PubMed=8670861;

RA Breeze A.L., Kara B.V., Barratt D.G., Anderson M., Smith J.C.,
 RA Lute R.W., Best J.R., Carltide S.A.;
 RT "Structure of a specific peptide complex of the carboxy-terminal SH2
 RT domain from the p85 alpha subunit of phosphatidylinositol 3-kinase.";
 RL EMO J. 15:3579-3589(1996).
 RN [8]
 RP VARIANT ILE-326.
 RX MEDLINE=97184306; PubMed=9032108;
 RA Hansen T., Andersen C.B., Echwald S.M., Urhammer S.A., Clausen J.O.,
 RA Vestergaard H., Owens D., Hansen L., Pedersen O.;
 RT "Identification of a common amino acid polymorphism in the p85alpha
 RT regulatory subunit of phosphatidylinositol 3-kinase: effects on
 RT glucose disappearance constant, glucose effectiveness, and the
 RT insulin sensitivity index.";
 RL Diabetes 46:494-501(1997).
 RN [9]
 RP VARIANT INSULIN RESISTANCE GLN-409, AND VARIANT ILE-326.
 RX MEDLINE=20230645; PubMed=10768093;
 RA Baynes K.C.R., Beeton C.A., Panayotou G., Stein R., Soos M.,
 RA Hansen T., Simpson H., O'Rahilly S., Shepherd P.R., Whitehead J.P.;
 RT "Natural variants of human p85 alpha phosphoinositide 3-kinase in
 RT severe insulin resistance: a novel variant with impaired
 RT insulin-stimulated lipid kinase activity.";
 RL Diabetologia 43:321-331(2000).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE PILO CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -1- SUBUNIT: HETERODIMER OF A PILO (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- DISEASE: DEFECTS IN PIK3R1 ARE A CAUSE OF SEVERE INSULIN
 CC RESISTANCE.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- DATABASE: NAME=PI3K; NOTE=PI3K 1:6-12(2000);
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1773542685.g.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M61906; -; NOT_ANNOTATED_CDS.
 DR PIR: A58748; A38748.
 DR PDB: 1PKS; 31-MAY-94.
 DR PDB: 1PKT; 31-MAY-94.
 DR PDB: 1PHT; 07-DEC-95.
 DR PDB: 1PBW; 12-MAR-97.
 DR PDB: 1PIC; 17-SEP-97.
 DR PDB: 1AON; 25-FEB-98.
 DR PDB: 1AZG; 25-FEB-98.
 DR Genew: HGNC:8979; PIK3R1.
 DR MIM: 171833; -.
 DR InterPro: IPR001720; PI3Kkinase_p85.
 DR InterPro: IPR000198; RHOGAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RHOGAP; 1.
 DR PRINTS: PR00678; PI3KINASEP85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODom: PD000066; SH3; 1.
 DR PRODom: PD000093; SH2; 2.
 DR SMART: SMO0324; RHOGAP; 1.
 DR SMART: SMO0252; SH2; 2.
 DR SMART: SMO0326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.

DR PROSITE; PS50002; SH3; 1.
 KW SH3 domain; SH2 domain; Repeat; 3D-structure; Polymorphism;
 KM Disease mutation..
 FT DOMAIN 3 79 SH3..
 FT DOMAIN 333 428 SH2 1..
 FT DOMAIN 624 718 SH2 2..
 FT VARIANT 326 326 M->I..
 FT VARIANT 409 409 /FTID-VAR_010023..
 FT VARIANT 409 409 R->Q (IN SEVERE INSULIN RESISTANCE;
 FT REDUCTION OF INSULIN-STIMULATED
 FT ACTIVITY).
 FT FTID-VAR_010024..
 SQ SEQUENCE 724 AA; 83597 MW; 5D9493C0431C760D CRC64;
 Query Match 7.3%; Score 146.5; DB 1; Length 724;
 Best Local Similarity 24.7%; Pred. No. 0.0082;
 Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;
 QY 176 PEVQGMPSQISLRDSEVLEAEKYPHNQ-----RKESTHLENTQDEI 220
 DB 211 PEVQSESEYIDL-LKILRSPSIHOYWLTLQYLLKHFKLSCSSKNLNAVLSSEIF 268
 QY 221 -PLAISSSFTTSMH-----SVQNDHRCGMQPCSPQRCPPASCSPHENILPY 268
 DB 269 SPMLEFSAASDNTENLKIYELLSTENNERQ-----PAPALPP----- 309
 QY 269 KYTSRPPPEPK-----BSDRKDVOHNEWYIGEYSROAVEAFMKENKGSFLVRCSTK 322
 DB 310 -----KPKKPTTVANNNGNNNNKSLQNAEMYMGDISREVENEK-LRDTADGTFVLVDASTK 363
 QY 323 SKEEPPVAVFYENKVVYVK-IRFLRNQOFALGTGLGDEKDSVEDIEHYKN 376
 DB 364 MHGD-YTLTL--RKGNNKLIKIFHRDGYKGFSDPL-----TFSSVELINHYRN 410
 RESULT 5
 P85A_BOVIN STANDARD; PRT; 724 AA.
 AC P23727;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
 DE p85-alpha subunit) (Ptdins-3-kinase p85-alpha) (PI3K).
 GN PIK3R1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91191567; PubMed=1707345;
 RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-larrea F., Panayotou G.,
 RA Thompson A., Dhand R., Hsuan J., Yotey N., Smith A.D., Morgan S.J.,
 RA Courtneidge S.A., Parker P.J., Waterfield M.D.;
 RT "Characterization of two 85 kd proteins that associate with receptor
 RT tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase.";
 RL Cell 65:91-104(1991).
 RN [2]
 RP CIRCULAR DICHROISM AND FLUORESCENCE SPECTROSCOPY.
 RX MEDLINE=93049176; PubMed=1330535;
 RA Panayotou G., Bax B., Gout I., Federwisch M., Wrzblowski B., Dhand R.,
 RA Fry M.J., Blundell T.L., Wollmer A., Waterfield M.D.;
 RT "Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
 RT domain with a PDGF receptor phosphorylation site: structural features
 RT and analysis of conformational changes.";
 RL EMO J. 11:4261-4272(1992).
 RN [3]
 RP STRUCTURE BY NMR OF 1-84.
 RX MEDLINE=93272320; PubMed=7684655;
 RA Booker G.W., Gout I., Downing A.K., Driscoll P.C., Boyd J.,
 RA Waterfield M.D., Campbell I.D.;

RT "Solution structure and ligand-binding site of the SH3 domain of the
 RT p85 alpha subunit of phosphatidylinositol 3-kinase.";
 RL Cell 73:813-822(1993).
 RN [4]
 RP STRUCTURE BY NMR OF 314-431.
 RA MEDLINE=92357146; PubMed=1323062;
 RA Booker G.W., Breeze A.L., Downing A.K., Panayotou G., Gout I.,
 RA Waterfield M.D., Campbell I.D.;
 RT "Structure of an SH2 domain of the p85 alpha subunit of
 RT phosphatidylinositol-3-OH kinase.";
 RL Nature 358:684-687(1992).
 RN [5]
 RP STRUCTURE BY NMR OF 321-434.
 RA MEDLINE=97110350; PubMed=8952511;
 RA Guenther U.L., Liu Y., Sanford D., Bachovich W.W., Schaffhausen B.;
 RT "NMR analysis of interactions of a phosphatidylinositol 3'-kinase SH2
 RT domain with phosphotyrosine peptides reveals interdependence of major
 RT binding sites.";
 RL Biochemistry 35:15570-15581(1996).
 RN [6]
 RP STRUCTURE BY NMR OF 614-724.
 RA MEDLINE=98173872; PubMed=9512716;
 RA Siegel G., Davis B., Kristensen S.M., Sankar A., Linacre J.,
 RA Stein R.C., Panayotou G., Waterfield M.D., Driscoll P.C.;
 RT "Solution structure of the C-terminal SH2 domain of the p85 alpha
 RT regulatory subunit of phosphoinositide 3-kinase.";
 RL J. Mol. Biol. 276:461-478(1998).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M61745; AAA79511.1; -
 DR PTR: A38749; A38749.
 DR PDB: 2PNA; 3I-JAN-94.
 DR PDB: 2PNB; 3I-JAN-94.
 DR PDB: 2PNI; 3I-OCT-93.
 DR PDB: 1PNJ; 3I-OCT-93.
 DR PDB: 1BFI; 25-FEB-98.
 DR PDB: 1BFI; 25-FEB-98.
 DR InterPro: IPR001720; PI3K_kinase_p85.
 DR InterPro: IPR00198; RhoGAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhoGAP; 1.
 DR PRINTS: PRO0678; PI3K_kinase_p85.
 DR PRINTS: PRO0401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhoGAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 KW SH3 domain; SH2 domain; Repeat; 3D-structure.
 FT DOMAIN 3 79
 FT SH3.
 FT DOMAIN 129 273 GAP DOMAIN.

FT DOMAIN 333 428 SH2 1.
 FT DOMAIN 624 718 SH2 2.
 FT HELIX 340 347
 FT TURN 348 348
 FT STRAND 354 359
 FT STRAND 368 373
 FT STRAND 378 382
 FT STRAND 384 384
 FT STRAND 391 391
 FT STRAND 401 409
 FT HELIX 413 413
 FT STRAND 418 419
 FT TURN 422 422
 FT STRAND 427 427
 FT STRAND 427 427
 SQ SEQUENCE 724 AA; 83497 MM; EDDF6E754BBE7321 CMC64;
 Query Match 7.2%; Score 146; DB 1; Length 724;
 Best Local Similarity 22.4%; Pred. No. 0.0088;
 Matches 77; Conservative 49; Mismatches 103; Indels 114; Gaps 16;
 Oy 136 RKNKTPLEPPRLITLPKKYQPLPPEPSSR-----PLSQRHRTPEV----- 178
 Db 80 KKSPPTEKPPRPRLPVAPGSKTEADSEQASTLPDLAQFAPDPVAPPLIKLYEAI 139
 Oy 179 --QGMPQSIQR-----DLSEVLEAEKVP-----HNQRKESTHLENTQEI 220
 Db 140 EKKGLECSTLYRTQSSSNPAELRQLDDCDTASLDLMDVHVLADAFKRYLLDLPN-FVI 198
 Oy 221 PLATSSSFT----- 236
 Db 199 PVAVSSELSLAPEVQSSSEYIQLKKLRSPSPHQYWLTLQYLLKHFKLQSSKNL 258
 Oy 227 ONRHRGMOPQRCORPASCSPH-ENILP-----YKTSR-----PPPKR--- 280
 Db 259 LNA--RVISELFPILFPFPASSENTHLIKITITLSTENENQPPALDPKPKPTT 316
 Oy 281 -----SDRKDVQHNEMYIGEYSRQAVEAFMKENKDSFLVDCSTKSKEPYLAIFY 334
 Db 317 VANNGMNNMNSLDQAEWYWGDISREVENK-LRDTAGDTFLVROASTKMHQD-YLLTL-- 372
 Oy 335 EKKYVNVK-IRPLENQOPALCTGIRGDEKPDVSDEITIEHKN 376
 Db 373 -RKGNMNLKIFHRDGRYGFSDPL---TFNSVVELINHYKN 410
 RESULT 6
 P85A_RAT STANDARD; PRT: 724 AA.
 ID P85A_RAT
 AC Q63787; Q63790; P70544; O55085;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol (PI)-3-kinase regulatory alpha subunit (PI3-kinase
 DE p85-alpha subunit) (Ptdins-3-kinase p85-alpha) (PI3K).
 GN PI3K1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS P85-ALPHA AND P55-ALPHA).
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=96214979; PubMed=6621382;
 RA Inukai K., Anai M., van Breda E., Hosaka T., Katagiri H., Funaki M.,
 RA Fukushima Y., Ogihara T., Yazaki Y., Kikuchi M., Oka Y., Asano T.;
 RT "A novel 55-kDa regulatory subunit for phosphatidylinositol 3-kinase
 RT of the p85alpha gene.";
 RL J. Biol. Chem. 271:5317-5320(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).
 RC TISSUE=Liver;
 RX MEDLINE=97218222; PubMed=9065454;

RA	Inukai K,Funaki M., Ogihara T., Katagiri H., Kanda A., Anal M.,
RA	Fukushima Y., Hosaka T., Suzuki M., Shin B., Takata K., Yazaki Y.,
RA	Kikuchi M., Oka Y., Asano T.;
RT	"p5alpha gene generates three isoforms of regulatory subunit for
RT	phosphatidylinositol 3-kinase (PI 3-Kinase), p50alpha, p55alpha, and
RT	p85alpha, with different PI 3-kinase activity elevating responses to
RT	insulin.";
RL	J. Biol. Chem. 272:7873-7882(1997).
RN	[3]
RE	SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).
RC	TISSUE=Liver;
RX	MEDLINE=9707966; PubMed=8921377;
RA	Furman D.A., Cantley L.C., Carpenter C.L.;
RA	"Structural organization and alternative splicing of the murine
RT	phosphoinositide 3-kinase p85 alpha gene.";
RT	Genomics 37:113-121(1996).
CC	-1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC	KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC	THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC	NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
CC	AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES
CC	-1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC	SUBUNTS.
CC	-1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: P85-ALPHA (SHOWN HERE), P55-
CC	ALPHA AND P50-ALPHA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	-1- TISSUE SPECIFICITY: THE P85-ALPHA ISOFORM IS WIDELY EXPRESSED.
CC	EXPRESSION OF THE P55-ALPHA ISOFORM IS HIGHEST IN BRAIN AND
CC	SKELETAL MUSCLE. THE P50-ALPHA ISOFORM IS ABUNDANT IN LIVER WITH
CC	LOWER LEVELS IN BRAIN AND MUSCLE.
CC	-1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC	-1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@sib-ch.ch).
CC	-----
DR	EMBL; D64045; BAA18932.1; -;
DR	EMBL; D64048; BAA18933.1; -;
DR	EMBL; U50412; AAC52846.1; -;
DR	EMBL; D78486; BAA24426.1; -;
DR	HSSP; P23727; IABT.
DR	InterPro: IPRO01720; PI3Kinase_P85.
DR	InterPro: IPRO00198; RhogAP.
DR	InterPro: IPRO00980; SH2.
DR	InterPro: IPRO01452; SH3.
DR	Pfam: PF000017; SH2; 2.
DR	Pfam: PF000018; SH3; 1.
DR	Pfam: PF00620; RhogAP; 1.
DR	PRINTS; PR00678; PI3KINASEP85.
DR	PRINTS; PR00401; SH2DOMAIN.
DR	ProDom; PD000066; SH3; 1.
DR	ProDom; PD000093; SH2; 1.
DR	SMART; SM00324; RhogAP; 1.
DR	SMART; SM00252; SH2; 2.
DR	SMART; SM00326; SH3; 1.
DR	PROSITE; PS50001; SH2; 2.
DR	PROSITE; PS50002; SH3; 1.
KW	SH3 domain; SH2 domain; Repeat; Alternative splicing.
FT	DOMAIN 3 79
FT	DOMAIN 333 428
FT	DOMAIN 624 718
FT	VARSPLIC 1 270
FT	VARSPLIC 271 304
FT	VARSPPLIC 1 300
FT	VARSPPLIC 301 306
FT	SEQUENCE 724 AA; 83531 MW; 95C65CF612873B84 CRC64;
FT	MISSING (IN ISOFORM P50-ALPHA).
FT	ROPAIPA -> MNHNOT (IN ISOFORM P50-ALPHA).
FT	TVMTMEDLEDCARTDINCGLDMFYTEMOP (IN
FT	ISOFORM P55-ALPHA)
FT	MISSING (IN ISOFORM P50-ALPHA).
FT	ROPAIPA -> MNHNOT (IN ISOFORM P50-ALPHA).
FT	SEQUENCE 724 AA; 83531 MW; 95C65CF612873B84 CRC64;

```

Query Match Summary      7.2%: Score 145; DB 1; Length 724;
Best Local Similarity    23.4%; Pred No. 0.01;
Matches    93; Conservative 48; Mismatches 155; Indels 102; Gaps 18;

QY      7 PKNRSPRINSATGOYQRNNKRPLLDMERNFAAVLDGAKGSHDDDDPELRMEETQSIK 66
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      87 PKRP RP RP RP V A G S --- SKTEADTEQGVLTLPDLAEQFAPPDVA P L L --- IK 134
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      67 ILPARIKSEVADTHYKVAADTPRLDTRTISIGOTWNTQTRLEAVDKP-ISRDP 125
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      135 LLEAIKKKLE-CSTLYRQSSNPRL------RQLDCCDPVSDLDVF 177
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      126 SONIKDASVRNNKIPLP RP RP LTP L P K K Q Y L P R P E S S R P L S O R H T F P V Q C M P S O I 185
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      178 DEHVLADA-FKRYLADLP RP ---VIRPAVYVEMMSLNO-----EVSSEDTI 220
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      186 SLRDLSLEVLEAEKVP R H N O-----RKPESTHLENNQNTQETPLAIVSSSFTT 231
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      221 QL-LKKLIRSPNIPHOYVLTQYLTKHFKLSQASKKLNLNARALSEI---FSVLPFR 275
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      232 SMHSYQNNRHR-----GGMQPSQRCQRPASCSPHENILPRKYTSMKRP P P K 279
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      276 PAASSDNTLHKAVELLISAEMSERQRPALPRP RP K T S I A N N S M----- 322
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      280 RSDRQDVQHNEMYIGEYSHQAAVEAFMKENKDGSEFLVRDCSTSKKEEYPVLAVFENKRY 339
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      323 -NNNSNLQDAEAWYMGDISREEVNEK-LRDTAGTGTEFLVRASDTRKMHGQ-VTLTL---RRGG 376
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      340 NVK-IRFLERNQOFALGTGLRDEKDFSDVEDIIEHKYK 376
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      377 NNKLKIEHRDGKYGSPDPL---TFNSVSELTINHYRN 410
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

ID	BRK_MOUSE	STANDARD:	PRT:	659 AA.
AC	P35991: 061365;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	tyrosine-protein kinase BTK (EC 2.7.1.112) (Bruton's tyrosine kinase)			
DE	(Agamaglobulinemia tyrosine kinase) (ATK) (B cell progenitor kinase) (BPK) (kinase EMB).			
GN	BTK OR BPK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=93145329; PubMed=8425221;			
RA	Tsukada S., Saffrin D.C., Rawlings D.J., Parolini O., Allen R.C.,			
RA	Kliskak I., Sparkes R.S., Kubagawa H., Mohandas T., Quan S.,			
RA	Belmont J.W., Cooper M.D., Conley M.E., Witte O.N.;			
RT	"Deficient expression of a B cell cytoplasmic tyrosine kinase in			
RL	human X-linked agamaglobulinemia.";			
RL	Cell 72:279-290(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93236578; PubMed=8476425;			
RA	Yamada N., Kawakami Y., Kimura H., Fukumachi H., Baler G.,			
RA	Altman A., Kato T., Inagaki Y., Kawakami T.;			
RT	"Structure and expression of novel protein-tyrosine kinases, Emb and			
RL	Emt, in hematopoietic cells.";			
RL	Biochem. Biophys. Res. Commun. 192:231-240(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95081608; PubMed=7989760;			
RA	Sideras P., Mueller S., Shuels H., Jin H., Khan W.N., Nilsson L.,			
RA	Parkinson E., Thomas J.D., Branden L., Larsson I., Paul W.E.,			
RA	Rosen F.S., Alt F.W., Vetrici D., Smith C.I.E., Xanthopoulos K.G.;			
RT	"Genomic organization of mouse and human Bruton's agamaglobulinemia			

RT tyrosine kinase (Btk) loci.";
 RL J. Immunol. 153:5607-5617(1994).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN:EC129;
 RX MEDLINE=95352959; PubMed=7626884;
 RA Oeljen J.C., Liu X., Lu J., Allen R.C., Muzny D.M., Belmont J.W.,
 RA Gibbs R.A.;
 RT "Sixty-nine kilobases of contiguous human genomic sequence containing
 RT the alpha-galactosidase A and Bruton's tyrosine kinase loci.";
 RL Mamm. Genome 6:334-338(1995).
 [5]
 RP VARIANT XID CYS-28.
 RX MEDLINE=93324903; PubMed=8332901;
 RA Rawlings D.J., Saffran D.C., Tsukada S., Largaespada D.A.,
 RA Grimaldi J.C., Cohen L., Mohr R.N., Bazan J.F., Howard M.,
 RA Copeland N.G., Jenkins N.A., White O.N.;
 RT "Mutation of unique region of Bruton's tyrosine kinase in
 RT immunodeficient xid mice.";
 RL Science 261:358-361(1993).
 CC -1- FUNCTION: PLAYS A CRUCIAL ROLE IN B-CELL ONTOGENY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -1- PTM: Autophosphorylated on Tyr-223 and Tyr-551. The tyrosine
 CC phosphorylation of Tyr-223 may create a docking site for a SH2
 CC containing protein.
 CC -1- DISEASE: DEFECTS IN BTK ARE THE CAUSE OF MURINE X-LINKED
 CC IMMUNODEFICIENCY (XID).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: L08967; AAA37316.1; -;
 DR EMBL: L10627; -; NOT ANNOTATED_CDS.
 DR EMBL: L29788; AAA66943.1; -;
 DR EMBL: U58105; AAB47246.1; -;
 DR HSSP: Q06187; IBMN.
 DR MGD: MGI:88216; Btk.
 DR InterPro: IPR001562; BTK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00779; BTK; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRINTS: PR00402; TECTKINASE.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00107; BTK; 1.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 1.
 DR SMART: SM00219; TyrcK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 DR Transferrase; Tyrosine-protein kinase; Phosphorylation;
 KW ATP-binding; SH3 domain; SH2 domain; Disease mutation;
 FT DOMAIN 3 133
 FT DOMAIN 214 274
 FT DOMAIN 281 377
 FT DOMAIN 402 655
 FT NP_BIND 408 416
 FT BINDING 450 430
 FT ACT_SITE 521 521
 FT MOD_RES 223 223
 FT MOD_RES 551 551
 FT VARIANT 28 28
 FT MUTAGEN 41 41
 FT MUTAGEN 223 223
 FT MUTAGEN 430 430
 FT CONFLICT 67 67
 FT CONFLICT 123 123
 FT CONFLICT 197 197
 FT CONFLICT 450 450
 SQ SEQUENCE 659 AA; 76437 MW; E502B798BC36E223 CRC64;
 Query Match 7.1%; Score 143; DB 1; Length 659;
 Best Local Similarity 23.4%; Pred. No. 0.013; Indels 98; Gaps 11;
 Matches 60; Conservative 27; Mismatches 71;
 QY 125 RSONINGDASVRKNIPLPPRPRLTLPRKYOPLPPEPSSRPPLSQRRHTEPVQGMPSQ 184
 Db 171 RNSGLKPGSHRRKTRKPLPTPEEDQILKK--PLPEPTA-----Ap 210
 QY 185 ISRLDSEVLEA-EKYPHNO-----RKPESTHLENQOEPLAISSSFTTNSVON 238
 Db 211 ISTTELRKVALYDYVPMNANDQLKGEYFLTESN--LFW-----WRA 254
 QY 239 RDRGGMQSCSPQRCPPASCSPHENILPYKYTSWMPRPKRSDRKDVONHMYIGEYSR 298
 Db 255 RDKNG-----QEGYIPSNYIT--EAEDSIENYEMYSKHMTR 288
 QY 299 QAVEAFMKENKDGSLVNDCKSTRKKEPYVLAIFYENKYNKIRFLERNOQFALGTGL 358
 Db 289 SQAEQLKQEGKGEFTIVDSSKAGK-----YTVSV-----FASTG- 325
 QY 359 RGDKEFDSVEDITEHY 374
 Db 326 -----EPQGVLRHY 334
 RESULT 8
 3BP2_MOUSE
 ID 3BP2_MOUSE STANDARD; PRT; 559 AA.
 AC 006649;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SH3 domain-binding protein 2 (3BP-2).
 GN SH3BP2 OR 3BP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93174278; PubMed=8438166;
 RA Ren R., Mayer B.J., Cicchetti P., Baltimore D.;
 RT "Identification of a ten-amino acid proline-rich SH3 binding site.";
 RL Science 259:1157-1161(1993).
 CC -1- FUNCTION: BINDS DIFFERENTIALLY TO THE SH3 DOMAINS OF CERTAIN
 CC PROTEINS OF SIGNAL TRANSDUCTION PATHWAYS. BINDS TO
 CC PHOSPHATIDYLINOSITOLS; LINKING THE HEMOPOIETIC TYROSINE KINASE

FES TO THE CYTOPLASMIC MEMBRANE IN A PHOSPHORYLATION DEPENDENT MECHANISM (BY SIMILARITY).

-1- SIMILARITY: CONTAINS 1 PH DOMAIN.

-1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; L14543; AAA37121.1; .

DR MGD; MG1:1346349; Sh3bp2.

DR InterPro: IPR001849; PH.

DR InterPro: IPR000980; SH2.

DR Pfam: PF00017; SH2; 1.

DR Pfam: PF00189; PH; 1.

DR ProDom: PD000093; SH2; 1.

DR SMART; SM00233; PH; 1.

DR SMART; SM00252; SH2; 1.

DR PROSITE; PS50001; SH2; 1.

DR PROSITE; PS50003; PH DOMAIN; 1.

DR SH2 domain; SH3-binding.

KW DOMAIN 26 130 PH.

FT SITE 201 210 SH3-BINDING.

FT DOMAIN 205 212 POLY-PRO.

FT DOMAIN 236 240 POLY-PRO.

FT FT 455 553 SH2.

SEQUENCE 559 AA; 62208 MW; EDPELF11B259646E CRC64;

Query Match 6.9%; Score 140; DB 1; Length 559;

Best Local Similarity 19.7%; Pred. No. 0.017;

Matches 93; Conservative 54; Mismatches 156; Indels 168; Gaps 18;

QY 56 LRHEETWQSKILPAPRIKSEYADTHYFKVAND-----TPPLD 95

DB 82 MRAAEETSNKVPFKIHSKRHRWFESASEDERKSMARVREIGHFHEKKELD 141

QY 96 TRTSISIGPTWNTQRLFEVDKPRISHDVSONIKGDASYRKNKIP--LPPRP- 147

DB 142 TSDSSS-----DTDSYGVAVRPIDSLSSYMDNEDYHEDEDSYLEPDSFGPKLE 195

QY 148 -LITLPEKTIQPLPPEPSSRPPLS--QRHTF-----PEVQMPQSQISL-RDL 190

DB 196 DALTYPPAYPP-PPVPPKPAFSDLPRAHSFTSKSPPLPPPPKRGLPDNGSAPEDA 254

QY 191 SEVLEAEKVPHNORKEPSTHLLNOMTQELPLAIS----- 225

DB 255 KDLGLRVRERGLRVPATPRKMSDPKSNVPTVNLKRNKCFRDSVNPGLPWTGHTS 314

QY 226 --SSSTTSHNSVQNRD-----HRGGMQ-----CSFORQO--- 254

DB 315 SVSSSTTMVAATSRNCDKLKSFLLSSRGPTSPPPYPAKPKFKLIAEPSTREAKFA 374

QY 255 --PPASCSPEHNTLPYKYSWRPPFKRSRDKVQH----- 288

DB 375 PVPVPVAPRPVQKMPBEAVBRAPVLPREPNTPLPHLQSRPDPGQSPRGSEFKARQPSQ 434

QY 289 -----NEWITGEYSKQAVDEAFMK-----ENKQGSFLVRCQSTKSKTE 335

DB 435 ADTGEDESDDEYKVKPLPSNVFNTTESCEVERLFKATDPGRGPODGLYXIRNSTSKGK 494

QY 326 EPYVLAVFEY--NKVYNVKIRFLERNQOFALGTLGDEKDEKFSVEDIIEHY 374

DB 495 ----VLVVMDESSKKVKRYRL--FEKDSKFT-----LGEVLFASVGSKEVHY 536

DT 01-JUN-1994 (Rel. 29, Created)

DT 15-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tyrosine-protein kinase BTK (EC 2.7.1.112) (Bcrton's tyrosine kinase)

DE (Agammaglobulinemia tyrosine kinase) (ATK) (B cell progenitor kinase) (BPK).

GN BTK OR ATK OR AGMX1 OR BPK.

OS Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID=9606;

OX [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=93140868; PubMed=8380905;

RA Vetric D., Vorechovsky I., Sideras P., Holland J., Davies A., RA Flinter F., Hammarstrom L., Kinnon C., Levinsky R.J., Bobrow M., RA Smith C.I.E., Bentley D.R.;

RT "The gene involved in X-linked agammaglobulinemia is a member of the RT src family of protein-tyrosine kinases.";

RL Nature 361:226-233(1993).

RN [2]

RP ERRATUM.

RA Vetric D., Vorechovsky I., Sideras P., Holland J., Davies A., RA Flinter F., Hammarstrom L., Kinnon C., Levinsky R.J., Bobrow M., RA Smith C.I.E., Bentley D.R.;

RL Nature 364:362-362(1993).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Blood;

RX MEDLINE=943377492; PubMed=8090769;

RA Ohta Y., Haire R.N., Litman R.T., Fu S.M., Nelson R.P., Kratz J., RA Kornfeld S.J., La Morena M., Good R.A., Litman G.W.;

RT "Genomic organization and structure of Bcrton agammaglobulinemia RT tyrosine kinase: localization of mutations associated with varied RT clinical presentations and course in X chromosome-linked RT agammaglobulinemia.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:9062-9066(1994).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=95012452; PubMed=7927535;

RA Rohrer J., Parolint O., Belmont J.W., Conley M.E.;

RT "The genomic structure of human BTK, the defective gene in X-linked RT agammaglobulinemia.";

RL Immunogenetics 40:319-324(1994).

RN [5]

RP SEQUENCE FROM N.A.

RX MEDLINE=95352959; PubMed=7626884;

RA Oeltjen J.C., Liu X., Lu J., Allen R.C., Muzny D.M., Belmont J.W., RA Gibbs R.A.;

RT "Sixty-nine kilobases of contiguous human genomic sequence containing RT the alpha-galactosidase A and Bcrton's tyrosine kinase loci.";

RL Mamm. Genome 6:334-338(1995).

RN [6]

RP SEQUENCE FROM N.A.

RA Oeltjen J.C., Malley T.M., Muzny D.M., Miller W., Gibbs R.A., RA Belmont J.W.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

RN [7]

RP SEQUENCE OF 1-442 FROM N.A.

RX MEDLINE=93145329; PubMed=8425221;

RA Tsukada S., Saffian D.C., Rawlings D.J., Parolint O., Allen R.C., RA Kilsak I., Sparkes R.S., Kubagawa H., Mohandas T., Quan S., RA Belmont J.W., Cooper M.D., Conley M.E., Witte O.N.;

RT "Deficient expression of a B cell cytoplasmic tyrosine kinase in RT human X-linked agammaglobulinemia.";

RL Cell 72:279-290(1993).

RN [8]

RP PHOSPHORYLATION OF GTF2I, AND MUTAGENESIS.

RX MEDLINE=97165069; PubMed=9012831;

RA Yang W., Desiderio S.;

RT "BAP-135, a target for Bcrton's tyrosine kinase in response to B cell RT receptor engagement.";

RL Proc. Natl. Acad. Sci. U.S.A. 94:604-609(1997).

[9] X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-170.
 RP MEDLINE-97361821; PubMed-9218782;
 RA Hyonenen M., Saraste M.;
 RT "Structure of the PH domain and Btk motif from Bruton's tyrosine
 kinase: molecular explanations for X-linked agammaglobulinemia.";
 RL EMBO J. 16:3396-3404(1997).
 RN
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-170.
 RX MEDLINE-99216540; PubMed-10196129;
 RA Baraldi E., Carrugo K.D., Hyonenen M., Surdo P.L., Riley A.M.,
 RA Potter B.V.L., O'Brien R., Laddbury J.E., Saraste M.;
 RT "Structure of the PH domain from Bruton's tyrosine kinase in complex
 with inositol 1,3,4,5-tetrakisphosphate.";
 RL Structure 7:449-460(1999).
 RN
 RP STRUCTURE BY NMR OF 209-275.
 RX MEDLINE-96153176; PubMed-9485443;
 RA Hansson H., Mattsson P.T., Allard P., Haapaniemi P., Vihtonen M.,
 RA Smith C.I.E., Haerd T.;
 RT "Solution structure of the SH3 domain from Bruton's tyrosine kinase.";
 RL Biochemistry 37:2912-2924(1998).
 RN
 RP REVIEW ON XLA VARIANTS.
 RX MEDLINE-96174621; PubMed-8594569;
 RA Vihtonen M., Iwata T., Kinnon C., Kwan S.-P., Ochs H.D.,
 RA Vorechovsky I., Smith C.I.E.;
 RT "BTKbase, mutation database for X-linked agammaglobulinemia (XLA).";
 RL Nucleic Acids Res. 24:160-165(1996).
 RN
 RP REVIEW ON XLA VARIANTS.
 RX MEDLINE-97169387; PubMed-9016530;
 RA Vihtonen M., Belohradsky B.H., Haire R.N., Hollinski-Feder E.,
 RA Kwan S.-P., Lappalainen I., Lehtvaslahti H., Lester T., Meindl A.,
 RA Ochs H.D., Ollila J., Vorechovsky I., Weiss M., Smith C.I.E.;
 RT "BTKbase, mutation database for X-linked agammaglobulinemia (XLA).";
 RL Nucleic Acids Res. 25:166-171(1997).
 RN
 RP VARIANTS XLA TRP-288; GLY-307; ASP-607 AND S-V-F-S-T-R-103 INS.
 RX MEDLINE-94214481; PubMed-8162056;
 RA Bradley L.A.D., Sweetman A.K., Lovering R.C., Jones A.M., Morgan G.,
 RA Levinsky R.J., Kinnon C.;
 RT "Mutation detection in the X-linked agammaglobulinemia gene, BTK,
 using single strand conformation polymorphism analysis.";
 RL Hum. Mol. Genet. 3:79-83(1994).
 RN
 RP VARIANTS XLA HIS-28 AND TRP-288.
 RX MEDLINE-94214435; PubMed-8162018;
 RA de Weers M., Mensink R.G.J., Kraakman M.E.M., Schuurman R.K.B.,
 RA Hendriks R.W.;
 RT "Mutation analysis of the Bruton's tyrosine kinase gene in X-linked
 agammaglobulinemia: identification of a mutation which affects the
 same codon as is altered in immunodeficient xid mice.";
 RL Hum. Mol. Genet. 3:161-166(1994).
 RN
 RP SEQUENCE FROM N.A., AND VARIANTS XLA S-334; R-506; Q-520; W-562 AND
 K-630.
 RX MEDLINE-95152493; PubMed-7880320;
 RA Hagemann T.L., Chen Y., Rosen F.S., Kwan S.-P.;
 RT "Genomic organization of the Btk gene and exon scanning for mutations
 in patients with X-linked agammaglobulinemia.";
 RL Hum. Mol. Genet. 3:1743-1749(1994).
 RN
 RP VARIANTS XLA D-113; C-361; Q-520; P-542; W-562; K-630 AND P-652.
 RX MEDLINE-95152494; PubMed-7849697;
 RA Conley M.E., Fitch-Hilgenberg M.E., Cleveland J.L., Parolini O.,
 RA Rohrer J.;
 RT "Screening of genomic DNA to identify mutations in the gene for
 Bruton's tyrosine kinase.";
 RL Hum. Mol. Genet. 3:1751-1756(1994).
 RN
 RP VARIANTS XLA H-28; P-33; P-408; G-589; D-613 AND 260-Q--E-280 DEL.
 RX MEDLINE-95153222; PubMed-7849721;

RA Zhu Q., Zhang M., Winkelstein J., Chen S.-H., Ochs H.D.;
 RT "Unique mutations of Bruton's tyrosine kinase in fourteen unrelated
 X-linked agammaglobulinemia families.";
 RL Hum. Mol. Genet. 3:1899-1900(1994).
 RN
 RP VARIANTS XLA E-430; Q-520; Q-525; P-562; V-582; G-589; E-594 AND
 D-613.
 RX MEDLINE-95108046; PubMed-7809124;
 RA Vihtonen M., Vetrle D., Maniar H.S., Ochs H.D., Zhu Q., Vorechovsky I.,
 RA Webster A.D.B., Notarangelo L.D., Nilsson L., Sowaski J.M.,
 RA Smith C.I.E.;
 RT "Structural basis for chromosome X-linked agammaglobulinemia: a
 tyrosine kinase disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12803-12807(1994).
 RN
 RP VARIANT XLA PHE-64, AND CHARACTERIZATION OF OTHER XLA VARIANTS.
 RX MEDLINE-95151728; PubMed-7849006;
 RA Vihtonen M., Zvelebil J.J.M., Zhu Q., Brocimans R.A., Ochs H.D.,
 RA Zegers B.J.M., Nilsson L., Waterfield M.D., Smith C.I.E.;
 RT "Structural basis for pleckstrin homology domain mutations in
 X-linked agammaglobulinemia.";
 RL Biochemistry 34:1475-1481(1995).
 RN
 RP VARIANTS XLA S-25; W-288; M-370; V-509; P-525; K-526; W-562; V-582
 AND R-594.
 RX MEDLINE-95227177; PubMed-7711734;
 RA Vorechovsky I., Vihtonen M., de Saint Basile G., Honsora S.,
 RA Hammarstrom L., Mueller S., Nilsson L., Fischer A., Smith C.I.E.;
 RT "DNA-based mutation analysis of Bruton's tyrosine kinase gene in
 patients with X-linked agammaglobulinemia.";
 RL Hum. Mol. Genet. 4:51-58(1995).
 RN
 RP VARIANTS XLA LYS-567; LEU-587 AND HIS-641.
 RX MEDLINE-95359977; PubMed-7633420;
 RA Jin H., Webster A.D.B., Vihtonen M., Sideras P., Vorechovsky I.,
 RA Hammarstrom L., Bernatowska-Matuszkiewicz E., Smith C.I.E.,
 RA Bobrow M., Vetrle D.;
 RT "Identification of Btk mutations in 20 unrelated patients with
 X-linked agammaglobulinemia (XLA).";
 RL Hum. Mol. Genet. 4:693-700(1995).
 RN
 RP VARIANTS XLA PRO-33; GLN-520; CYS-641 AND GLY-302 DEL.
 RX MEDLINE-95359987; PubMed-7633429;
 RA Gaspar H.B., Bradley L.A.D., Katz F., Lovering R.C., Rolfman C.M.,
 RA Morgan G., Levinsky R.J., Kinnon C.;
 RT "Mutation analysis in Bruton's tyrosine kinase, the X-linked
 agammaglobulinemia gene, including identification of an insertional
 hotspot.";
 RL Hum. Mol. Genet. 4:755-757(1995).
 RN
 RP VARIANTS XLA ASN-429 AND ARG-477.
 RX MEDLINE-96177680; PubMed-8634718;
 RN
 RX
 Query Match 6.9%; Score 140; DB 1; Length 659;
 Best local Similarity 22.7%; Pred. No. 0.02;
 Matches 58; Conservative 29; Mismatches 71; Indels 98; Gaps 11;
 QY 125 RQNTINGDASVKKNKIPDPRLPILPKKYQLPPEPSSRRPLSQRHTFFVQGMPSQ 184
 DB 171 RNSGLKPGSSHRTKPLPTPEEDQILKK--PLPEPAA-----Ap 210
 QY 185 ISLRDSEVLEA-EKVPNHQ-----RKPESTHLENNQOEIPLAISSSFTSNHNVON 238
 DB 211 VTSSELKRVVALLDYPMNANDQLKGDGYFLLESN---LPW-----WRA 254
 QY 239 RDHGGMOQSPORCOPPASCSPHENILPYKYTSWBPPEPKRSRDRKDVQHNEMYIGEYSR 298
 DB 255 RDKNG-----OEGYIPSNVY-----EAEDSIEMYYSKHMTR 288
 QY 299 QAVEEAPFKENKGSFLVDCSTKSKPEPYVLAFTENKYNVKKIFLENNQOPALCTGL 358
 DB 289 SQAEOLLKQEGKGGFIVRDSKAGK-----YTVSV-----FAKSTG- 325

QY 359 RGDGKFDSDVEDIEHY 374
 Db 326 -----DPOGVIRHY 334

RESULT 10

CSW_DROME STANDARD; PRT: 845 AA.
 ID CSW_DROME
 AC P29349; Q24032; Q9W524; Q24033; Q9V3H1;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein-tyrosine phosphatase corkscrew (EC 3.1.3.48).
 GN CSM OR EG:BACN25G24.2 OR CG3954.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORM Y1229), FUNCTION, TISSUE SPECIFICITY, AND
 RP DEVELOPMENTAL STAGE.
 RC TISSUE-Embryo;
 RX MEDLINE=92346711; PubMed=1638629;
 RA Perkins L.A., Larsen I., Perrimon N.;
 RT "Corkscrew encodes a putative protein tyrosine phosphatase that
 RT functions to transduce the terminal signal from the receptor tyrosine
 RT kinase torso.";
 RL Cell 70:225-236(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS Y1229 AND 4A).
 RC STRAIN-DP CN BW;
 RA Melnick M.B., Melnick C.B., Larsen I., Perrimon N., Perkins L.A.;
 RT "The role of the Drosophila corkscrew protein as a transducer
 RT downstream of receptor tyrosine kinases is functionally conserved.";
 RT Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS Y1229; 2 AND 4A).
 RC STRAIN-Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherter S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista A.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fodor B.C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Idegawa C.,
 RA Jaitai M., Kalush F., Kapen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Moshfegh D.,
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murthy D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 4A AND Y1229).
 RC STRAIN-Oregon-R;
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Bartell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
 RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borokova D.,
 RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
 RA Papadogiannakis G., Spanos L., Cox S., Medueno E., de Pablo B.,
 RA Pappalini J., Peter A., Schoettler P., Werner M., Mourikoti F.,
 RA Betnert N., Dove G., Schaefer U., Jaekle H., Bucheton A.,
 RA Callister D.M., Campbell L.A., Darlamiotou A., Henderson N.S.,
 RA McMillan P.J., Sallies C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.,
 RT "From sequence to chromosome: the tip of the X chromosome of D.
 RT melanogaster.";
 RL Science 287:2220-2222(2000).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=97105827; PubMed=8948575;
 RA Perkins L.A., Johnson M.R., Melnick M.B., Perrimon N.;
 RT "The nonreceptor protein tyrosine phosphatase corkscrew functions in
 RT multiple receptor tyrosine kinase pathways in Drosophila.";
 RL Dev. Biol. 180:63-81(1996).
 CC -1- FUNCTION: REQUIRED IN ALL RECEPTOR TYROSINE KINASE SIGNALING
 CC PATHWAYS. FUNCTIONS DOWNSTREAM OF THE RECEPTOR TYROSINE KINASE
 CC TORSO, ACTING IN CONCERT WITH D-RAF VIA TAILLESS. ALSO FUNCTIONS
 CC DOWNSTREAM OF EGFR (EPIDERMAL GROWTH FACTOR RECEPTOR) AND BTL
 CC (PIRHOBLAST GROWTH FACTOR RECEPTOR). THE SH2 DOMAIN SUGGESTS THAT
 CC CSM EFFECTS ITS ROLE BY MEDIATING HETEROLOGIC PROTEIN
 CC INTERACTIONS. MATERNAALLY REQUIRED FOR NORMAL DETERMINATION OF CELL
 CC FATES AT THE TERMINI OF THE EMBRYO. REQUIRED FOR CELL FATE
 CC SPECIFICATION OF THE VENTRAL ECTODERM. IN THE DEVELOPING EMBRYONIC
 CC CNS AND FOR EMBRYONIC TRACHEAL CELL MIGRATION. FUNCTIONS DURING
 CC IMAGINAL DEVELOPMENT FOR PROPER FORMATION OF ADULT STRUCTURES SUCH
 CC AS EYES, ARISTAE, US WING VEIN AND THE TARSAL CLAW.
 CC CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 2 (SHOWN HERE), 4A AND Y1229;
 CC MAY BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED UNIFORMLY THROUGHOUT ALL TISSUES
 CC DURING EMBRYOGENESIS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT.
 CC WHICH SHARES NO HOMOLOGIES WITH OTHER PTASE PROTEINS. THIS PTASE
 CC INSERT IS REMINISCENT OF THE KINASE INSERT WITHIN THE KINASE
 CC CATALYTIC DOMAINS OF SEVERAL RECEPTOR TYROSINE KINASES.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. SUBCLAS THAT CONTAINS SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M94730; AAA28433.1; -
 CC EMBL: U19909; AAB02543.1; -
 CC EMBL: U19909; AAB02544.1; -
 CC EMBL: AE003423; AAF45724.1; -
 CC EMBL: AE003423; AAF45724.1; -
 CC EMBL: AE003423; AAF45724.1; -
 CC EMBL: AE003423; AAF45725.1; -
 CC EMBL: AL132797; CAB65870.1; -
 CC EMBL: AL132797; CAB65871.1; -

DR PIR: A43254; A43254.
 DR HSP: Q06124; 2SH2.
 DR FlyBase: FBgn0000382; CSW.
 DR InterPro: IPR003595; PTPC_motif.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR InterPro: IPR000242; TYR_PP.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00102; Y_phosphatase; 1.
 DR PRINTS: PR00700; PRTYHPHTASE.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD00093; SH2; 2.
 DR SMART: SM00194; PTPC; 1.
 DR SMART: SM00404; PTPC_motif; 1.
 DR SMART: SM00252; SH2; 2.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR Hydrolase; SH2 domain; Repeat; Developmental protein;
 KW Alternative splicing.
 FT DOMAIN 6 101 SH2 1.
 FT DOMAIN 111 205 -SH2 2.
 FT DOMAIN 227 645 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 289 444 PTPASE INSERT (CYS/SER-RICH).
 FT ACT_SITE 583 583 BY SIMILARITY.
 FT VARSPPLIC 1 110 MSSRRFPTTISGIAEAEKLOEGDGSFLARLSSNGAF
 TLYVRGNEVTHIKIIONNGDFDLGEGKFPILPTLYYM
 ENGELKERNGOAEIKOPLICAETTER -> MLENKLEK
 LSSSLGANNHLOEKQYNNNNNNNNNNNNNNNNNNNNQ
 KNFEYERAIQAHYSGKRRSEERSSGFKASKGRKAVTP
 PTEPEAEQPAKNCMTDELAOLIKGYAKGADARNDNR
 LQRRRPLSAQPSAASASTSTESLHRLTPQASYPATPT
 SWTATPPQPPAFAFGASCSNSTSLATMRVQLGHEYT (IN
 ISOFORM 4A).
 FT FT CMOQRPAPLPREKILKILNSPYFQONSKTFPRT -> A
 KFKNIPKDMIGLRPSPHAPLPPTPPPKT (IN
 ISOFORM Y1229).
 FT FT CONFLICT 749 749 G -> S (IN REF. 2).
 FT FT CONFLICT 815 815 P -> S (IN REF. 2).
 FT FT SEQUENCE 845 AA; 92975 MM; 2147F0F2576202CC CRC64;
 SO
 Query Match 6.8%; Score 136.5; DB 1; Length 845;
 Best Local Similarity 23.4%; Pred. No. 0.047;
 Matches 54; Conservative 39; Mismatches 67; Indels 71; Gaps 9;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RX MEDLINE=97446134; PubMed=9299232;
 RA Bell S.M., Shaw M., Joo Y.-S., Myers R.M., Knowles M.A.;
 RT "Identification and characterization of the human homologue of
 SH3BP2, an SH3 binding domain protein within a common region of
 deletion at 4p16.3 involved in bladder cancer."; Genomics 44:163-170(1997).
 RL Genomics 44:163-170(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS), AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=96403881; PubMed=9734812;
 RA Hadano S., Ishida Y., Ikeda J.-E.;
 RT "The primary structure and genomic organization of five novel
 transcripts located close to the Huntington's disease gene on human
 chromosome 4p16.3."; DNA Res. 5:177-186(1998).
 RL DNA Res. 5:177-186(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RC TISSUE=Tonsil;
 RA Gokemeijer J., Deligianidis K.E., Ligris K., Ernst T.J.;
 RT "3BP2 binds to phosphatidylinositols; linking the hemopoietic tyrosine
 kinase c-FES to the cytoplasmic membrane in a phosphorylation
 dependent mechanism."; Blood 88:473A-473A(1996).
 RL Blood 88:473A-473A(1996).
 RN [4]
 RP VARIANTS CRBM GLN-415; PRO-415; ARG-418; HIS-418; LEU-418; ARG-420 AND
 GLU-420.
 RX MEDLINE=21275962; PubMed=11381256;
 RA Ueki Y., Tiziani V., Santana C., Fukui N., Maulik C., Garfinkle J.,
 RA Nimmiya C., domarai S., Peters H., Hahel N., Rheo-Morris L.,
 RA Doss J.B., Kreiborg S., Olsen B.R., Reichenberger E.;
 RT "Mutations in the gene encoding c-Abi-binding protein SH3BP2 cause
 cherubism."; Nat. Genet. 28:125-126(2001).
 RL Nat. Genet. 28:125-126(2001).
 CC -1- FUNCTION: BINDS DIFFERENTIALLY TO THE SH3 DOMAINS OF CERTAIN
 PROTEINS OF SIGNAL TRANSDUCTION PATHWAYS. BINDS TO
 PHOSPHATIDYLINOSITOLS; LINKING THE HEMOPOIETIC TYROSINE KINASE
 FES TO THE CYTOPLASMIC MEMBRANE IN A PHOSPHORYLATION DEPENDENT
 MECHANISM.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
 short form; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in a variety of tissues including
 lung, liver, skeletal muscle, kidney and pancreas.
 CC -1- DISEASE: Defects in SH3BP2 are the cause of cherubism (CRBM), an
 autosomal dominant inherited syndrome. It is characterized by
 excessive bone degradation of the upper and lower jaws, which
 often begins around three years of age. It is followed by
 development of fibrous tissue masses, which causes a
 characteristic facial swelling.
 CC CHARACTERISTIC: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC EMBL: U56386; AAB72034.1; -
 DR EMBL: AB000462; BAA19119.1; -
 DR EMBL: AB000463; BAA19120.1; -
 DR EMBL: AF000936; AAB59973.1; -
 DR GenBank: HGNC:10825; SH3BP2.
 DR MIM: 602104; -
 DR MIM: 118400; -
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000980; SH2.

RL Gene 228:243-252(1999).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-Berkeley.
 RA MEDLINE-20196006; Pubmed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Planck C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheele F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Styrbas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: ADAPTER PROTEIN WHICH INTERACTS WITH C-TERMINAL PORTION
 CC OF MBC, HOMOLOG OF HUMAN DOCK180. MAY PLAY A ROLE IN CELLULAR
 CC PROCESSES THROUGHOUT DEVELOPMENT.
 CC -1- TISSUE SPECIFICITY: EMBRYONIC ZYGOTIC EXPRESSION IS SEEN IN
 CC INVAGINATING PRESUMPTIVE MESODERM AND ECTODERMALLY DERIVED TISSUES
 CC DURING GASTRULATION. AT STAGE 8, EXPRESSION IS ALSO SEEN IN
 CC ANTERIOR AND POSTERIOR MIDGUT AND CEPHALIC FURROW. BY STAGE 9,
 CC EXPRESSION IS HIGHEST IN VISCERAL MESODERM OF ANTERIOR AND
 CC POSTERIOR MIDGUT, VENTRAL NERVE CORD AND SOMATIC MESODERM.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY
 CC THROUGHOUT EMBRYOGENESIS. DECLINES DURING LARVAL STAGES AND
 CC REAPPEARS DURING PUPATION.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE003844; AA059362.1; -
 CC EMBL: AF112976; AAD28428.1; -
 CC HSSP: Q64010; ICKA.
 CC FlyBase: FBgn0024811; Crk.
 CC InterPro: IPR000980; SH2.
 CC InterPro: IPR001452; SH3.
 CC Pfam: PF00017; SH2; 1.

DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRODOM: PD000066; SH3; 1.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 2.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 1.
 KM SH2 domain; SH3 domain; Repeat; Developmental protein.
 FT DOMAIN 12 114 SH2.
 FT 117 165 SH3 1.
 FT DOMAIN 220 259 SH3 2.
 SQ SEQUENCE 271 AA; 31205 MW; D1BAFEA3150932DC CRC64;
 Query Match 6.38; Score 127; DB 1; Length 271;
 Best Local Similarity 35.18; Pred. No. 0.054;
 Matches 33; Conservative 14; Mismatches 37; Indels 10; Gaps 3;
 QY 281 SDRDVQHNENYIGEYSRQAVEAFMKENKDGSLVBCSTKSEPYVLAVFENKRVN 340
 DB 7 SDR-----NSWYFQPMNRQDQATEVLNMRERGVFLVDSNSIAD--YVLCVREDTVSN 59
 QY 341 VKIFLEPNOGALGTGLRDEKFDVSDIIEHY 374
 DB 60 YIINKVQODQDIYRI---GDQSFNDLPKLLTFY 90
 RESULT 15
 P85B_MOUSE STANDARD; PRT; 722 AA.
 AC P85B_MOUSE
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) (Ptdins-3-kinase p85-beta).
 GN PI3KR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIH Swiss;
 RX MEDLINE-98241181; Pubmed-9582025;
 RA Janssen J.M.G., Schleithoff L., Barttram C.R., Schulz A.S.;
 RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase
 RT p85beta subunit and HOMOXR8, a putative deubiquitinating enzyme";
 RL Oncogene 16:1767-1772(1998).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Y13569; CA73903.1; -
 CC HSSP: P23727; 2PNB.
 CC MGD: MGI:1098772; PI3kr2.
 CC InterPro: IPR001720; PI3kinase_P85.
 CC InterPro: IPR000198; RHOgap.

DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00017; SH2; 2.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00620; RhogAP; 1.
DR PRINTS: PR00678; PI3KINASEP85.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000066; SH3; 1.
DR PRODOM: PD000093; SH2; 2.
DR SMART: SM00324; RhogAP; 1.
DR SMART: SM00252; SH2; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50001; SH2; 2.
DR PROSITE: PS50002; SH3; 1.
DR SH3 domain; SH2 domain; Repeat.
KW DOMAIN 4 SH3.
FT DOMAIN 125 257 RHO-GAP.
FT DOMAIN 324 419 SH2 1.
FT DOMAIN 616 710 SH2 2.
SQ SEQUENCE 722 AA; 81251 MW; AB4D49AF30CEC567 CRC64;

Query Match

6.3%; Score 127; DB 1; Length 722;

Best Local Similarity 23.5%; Pred No 0.17; Mismatches 146; Indels 129; Gaps 28;

Matches 100; Conservative 50; Mismatches 146; Indels 129; Gaps 28;

OY 7 PNRSM-PRINSAT-----GQYRNNKPL---LDWERNFAAV 39
DB 51 PHNVGMPCGFENRTQRGDPCGYVEFLGPVALARPGPRPRPLPARPLDSSSESHI 110
OY 40 L-DGAKGSHDDYDDELMETWQSIKILPARPIKESYADTHYKAMDTPLDITRT 98
DB 111 LFDLAEGQSPDPAPPIL-----VKLYEA--IEQAEIDSECYSK---PELPA-TRT 155
OY 99 SISIGQPTWNTQRLERVNKPISRDRVRSQNIKGDAVSRKNKILPPEPRPLT---LPKKY 155
DB 156 DMSLSD-----LEQMDRTALYDA---VKG-----FLALPAAVTPEAAAEAY 195
OY 156 QPL-----PEPESSRPPLSQRHTFPEVOGMPQSILRLSEVLEAEKYVPHNQRKE-STH 210
DB 196 RALREVAGPVGLVLEPP-----TLPLHQALTLRFLQLHGLGRV--ARRAP---SPDTAVH 244
OY 211 LL-----ENONTQELP-LAISSSFSTSNHSVQNRDHRGGMQPCSPQRC 253
DB 245 ALASAFGLLLRIPPSGEGDSEVPDPFVILLERLYOEH--VERQDAAPALPPKPSKA 303
OY 254 QP-PASCSPHENILPYKXTSMRPPPKRSRDKVOHNEMVIGEYSROAVEEAFMKENKDG 312
DB 304 KPAFTALA-----NGGSP-----SLQDAEMTWGDISREYNER-LRDTPDG 344
OY 313 SELVDCSTKSEKPYVLAIFYENKVVYVK-IRFLERNQOAFALGTGLRDEKFDSEYEDIT 371
DB 345 TELVADASSKIGE-YTLTL--RKGNNKLKVFHRDGHYGFSEPL---TFCSVVELI 396
OY 372 EHYKN 376
DB 397 SHYRH 401

Search completed: April 21, 2003, 12:36:55
Job time : 12.3453 secs

